

O'Bry n, Barbara

From:
Sent:
To:
Subject:

Navarro, Albert
Monday, February 05, 2001 11:07 AM
O'Bryen, Barbara
09/001,737

Mark Navarro
1645
306-3225
8A15

Hello Barb,

Could I trouble you for a search and an interference search of:

SEQ ID NO: 7

Thanks

Mark

POINT OF CONTACT:
BARB O'BRYEN
TECH. INFORMATION SPECIALIST
STIC CM1 12C14 308-4291

BOB
2-16-01

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2001, 20:51:12 ; Search time 1971.82 seconds
(without alignments)
4311.034 Million cell updates/sec

Title: US-09-001-737-7

Perfect score: 1661
Sequence: 1 GAATTCGGCTTCATATGCA.....TGGCGGATACCGCAATTC 1661

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1118133 seqs, 2558875100 residues

Total number of hits satisfying chosen parameters: 2236266

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_cm:*
4: gb_ov:*
5: gb_ph:*
6: gb_pl1:*
7: gb_pl2:*
8: gb_pr1:*
9: gb_pr2:*
10: gb_pr3:*
11: gb_ro:*
12: gb_sy:*
13: gb_un:*
14: em_fun:*
15: em_hum1:*
16: em_hum2:*
17: em_in:*
18: em_cm:*
19: em_or:*
20: em_ov:*
21: em_pat:*
22: em_ph:*
23: em_pl:*
24: em_ro:*
25: em_sts:*
26: em_sy:*
27: em_un:*
28: em_vl:*
29: gb_ba3:*
30: gb_in1:*
31: gb_in2:*
32: gb_in3:*
33: gb_pl3:*
34: gb_pr4:*
35: em_ba1:*
36: em_ba2:*
37: em_htg1:*
38: em_htg2:*
39: em_htg3:*
40: em_htg4:*
41: em_htg5:*
42: em_htg6:*
43: em_htg7:*

44: em_htg8:*
45: em_htg9:*
46: em_htg10:*
47: em_hum3:*
48: em_hum4:*
49: em_hum5:*
50: em_hum6:*
51: gb_pr5:*
52: gb_pr6:*
53: gb_pr7:*
54: gb_htg1:*
55: gb_htg2:*
56: gb_htg3:*
57: gb_htg4:*
58: gb_htg5:*
59: gb_htg6:*
60: gb_htg7:*
61: gb_htg8:*
62: gb_htg9:*
63: gb_htg10:*
64: gb_htg11:*
65: gb_htg12:*
66: gb_htg13:*
67: gb_htg14:*
68: gb_htg15:*
69: gb_htg16:*
70: gb_htg17:*
71: gb_htg18:*
72: gb_htg19:*
73: gb_htg20:*
74: gb_htg21:*
75: gb_htg22:*
76: gb_htg23:*
77: gb_sts1:*
78: gb_sts2:*
79: gb_vl1:*
80: gb_vl2:*
81: gb_pat1:*
82: gb_pat2:*
83: em_htg0:*
84: em_htg24:*
85: gb_pr8:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1413.4	85.1	2683	29	SPGROELGN
2	1030.6	62.0	2401	1	AF117741
3	948.6	57.1	2267	2	LI6ROESL
4	840.4	50.6	2834	1	AB028452
5	838.8	50.5	2309	2	BACSGROESL
6	807.4	48.6	2400	2	BACSGROESL
7	807.4	48.6	2525	2	BACHSPA
8	807.4	48.6	17516	1	AB007637
9	807.4	48.6	213190	2	BSUB0004
10	799.2	48.1	2190	2	BACSGROEL
11	795.8	47.9	2047	1	AF259784
12	786.4	47.3	296950	2	AF001508
13	773	46.5	3128	1	AF010281
14	771	46.4	2753	1	AF214488
15	756.4	45.5	3885	1	AF269843
16	751.6	45.2	2187	29	SEU13618
17	739.8	44.5	2540	2	D55630
18	736.2	44.3	2675	1	AF031929
19	722.4	43.5	2712	29	STAHSP
20	704.2	42.4	4710	2	LI045241
21	688.8	41.5	3987	2	CPGROESL
					X89236 S.pyogenes
					AF117741 Streptococ
					X71132 L.lactis gr
					AB028452 Bacillus
					LI0132 Bacillus st
					M01132 Bacillus su
					M04965 B.subtilis
					AB007637 Bacillus
					Z99107 Bacillus su
					D10972 Bacillus su
					AF259784 Bacillus
					AF001508 Bacillus
					AF010281 Lactobaci
					AF214488 Lactobaci
					AF269843 Staphyloc
					LI3618 Staphylococ
					D55630 Bacillus sp
					AF031929 Lactobaci
					D14711 Staphylococ
					U45241 Lawsonia in
					X62914 C.perfringe

22	680.2	41.0	1940	1	AF080547	AF080547 Clostridi
23	674	40.6	2204	2	CTGROESEL	Z68137 C.thermocl
24	673.8	40.6	1623	1	AF159449	AF159449 Clostridi
25	657	39.6	1644	2	BH078514	U78514 Bartonella
26	656.4	39.5	2103	29	TBU56021	U56021 Thermoaer
27	652.6	39.3	3493	2	CLOGROESLA	M74572 Clostridium
28	652.2	39.3	1644	2	ROU78515	U78515 Bartonella
29	646.4	38.9	1627	1	AF014829	AF014829 Bartonella
30	645.6	38.9	1894	2	BAOB863A	M98257 Bartonella
31	644.6	38.8	1811	2	LPNHTPBG	M31918 Legionella
32	643.4	38.7	1788	2	PHA243594	U243594 Pseudofl
33	640.8	38.6	1635	2	BBCROELT	Z15160 B.bacillifo
34	636.6	38.3	5907	2	D85628	D85628 Buchnera ap
35	636.6	38.3	34699	1	AF008210	AF008210 Buchnera
36	630.6	38.0	282183	2	CJ11168X4	AL139077 Campyloba
37	629	37.9	2580	2	CJY13334	Y13334 Campylobact
38	615.8	37.1	2016	2	LPNHTPB	M91573 Legionella
39	615.8	37.1	2739	2	APSYMUL	X61150 Acyrthosiph
40	615.8	37.1	11763	29	U32736	U32736 Haemophilus
41	615.8	37.1	347550	2	AP001118	AP001118 Buchnera
42	614.2	37.0	6224	1	AF003957	AF003957 Myzus per
43	611.4	36.8	2596	2	FTGRO	X98853 F.tularensi
44	609.4	36.7	1522	1	AF014833	AF014833 Bartonella
45	603	36.3	2360	2	PM030165	U30165 Pasteurella

ALIGNMENTS

RESULT 1	SPGROELN	2683 bp	DNA	BCT	26-SEP-1997
LOCUS	S.pyogenes DNA for groEL gene.				
DEFINITION	X89236.1 GI:2462691				
VERSION	X89236.1				
KEYWORDS	groEL gene; heat shock protein 60 (GroEL) like protein.				
SOURCE	Streptococcus pyogenes.				
ORGANISM	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.				
REFERENCE	1 (bases 1 to 2683)				
AUTHORS	Podbielski, A.				
JOURNAL	Direct Submission				
REFERENCE	Submitted (29-JUN-1995) A. Podbielski, Institut f Med				
AUTHORS	Mikrobiologie, an der Rhein Westf. Techn. Hochschule, Aachen,				
JOURNAL	Pauwelsstr Klinikum, 52057 Aachen, FRG				
REFERENCE	2 (bases 1 to 2683)				
AUTHORS	Pohl, B., Podbielski, A. and Zarges, I.				
JOURNAL	Unpublished				
COMMENT	Related sequences M81132, M84965.				
FEATURES	Location/Qualifiers				
SOURCE	1. 2683				
gene	/organism="Streptococcus pyogenes"				
CDS	/strain="serotype M49"				
	/isolate="CS101"				
	/db_xref="taxon:1314"				
	1. 1446				
	/gene="groEL"				
	<1. 1446				
	/gene="groEL"				
	/codon_start=1				
	/transl_table=11				
	/product="heat shock protein"				
	/protein_id="CA61520.1"				
	/db_xref="gi:2462692"				
	/translation="DHFMNGKLYSEVASKNDIAGDSTTATVLTQAIYIEGKLV				
	TAGANPIGIRKGIETATATVEALKAVAQPVSGKAIAOVAVSSRSEVGEIIEAM				
	BRVNDGVITIEESRGMELEVEGMOFDRGYLSQYVTDNEKAVADLENFIIIMD				
	KVSNIQDILPLEEVLKTNRLITLADVDGEALPTVLNKIRGFNVAVKAKGFG				
	DRKRAMEDIALTCGVITEDLGLKDATATATGAOAKIIVDSIVYVSGSSSE				
	AIANRIALIKSLETTTSDFDREKIOEPLAKIAGVATKYKCAPEFLKKEKLIED				
	ALNATRAVEBSIVAGGCTALTIVIEKVALEBEDDATGNRIIVLRALDEEPVQIALN				

QY	201	GATCATTTTGAATAACATGGAGCAAAATTGGTGTCTGAAGTGGCTCTTAACCAATGAT	260	AGEGSVYIDKLKNSPAGTGFNNAATGENDMKTGIIIDPVKYTRSLAONASVASLIL
QY	1	GATCATTTTGAATAACATGGAGCAAAATTGGTGTCTGAAGTGGCTCTTAACCAATGAT	60	TTEAVVANKPEEPAAPAPAMPAGMDPDMGGF*
QY	261	ATTGCTGGTGGATGGGAGCACTACTGCAACGTTTGGACAAACCACTGGTCAATGAAGCA	320	repeal_region 1509. 1564
QY	61	ATTGCTGGTGGATGGGAGCACTACTGCAACGTTTGGACAAACCACTGGTCAATGAAGCA	120	/rpl_type=INVERTED
QY	321	CTAATAAATGTGACAGCAGGTGCTAATCCAAATTGATATCCGTGAGGCAATTAACAGCA	380	BASE COUNT 808 a 490 c 593 g 792 t
QY	121	CTAATAAATGTGACAGCAGGTGCTAATCCAAATTGATATCCGTGAGGCAATTAACAGCA	180	ORIGIN
QY	381	ACAGCAACAGCTGTGAAGCTTGAAGCCATTGCTCAACCTGTATCTGGCAAGAACT	440	Query Match 85.1%; Score 1413.4; DB 29; Length 2683;
QY	181	ACAGCAACAGCTGTGAAGCTTGAAGCCATTGCTCAACCTGTATCTGGCAAGAACT	240	Best Local Similarity 98.9%; Pred. No. 2.1e-257;
QY	441	ATTGCTGGTGGATGGGAGCACTACTGCAACGTTTGGACAAACCACTGGTCAATGAAGCA	500	Matches 1423; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY	241	ATTGCTGGTGGATGGGAGCACTACTGCAACGTTTGGACAAACCACTGGTCAATGAAGCA	300	
QY	501	GCTATGAGAGGTGTGGGCAACGATGGTGTGATACCATGCAAAATCTGAGGTATGAA	560	
QY	301	GCTATGAGAGGTGTGGGCAACGATGGTGTGATACCATGCAAAATCTGAGGTATGAA	360	
QY	561	ACAGAACTTGAAGTGTGAAGGATGCAATTTGACCGTGGTATCTGTCTCAATACATG	620	
QY	361	ACAGAACTTGAAGTGTGAAGGATGCAATTTGACCGTGGTATCTGTCTCAATACATG	420	
QY	621	GTCACAGACAAATGAAAAATGTTGACAGACCTTGAAGCCATTATCTTAACACAGAT	680	
QY	421	GTCACAGACAAATGAAAAATGTTGACAGACCTTGAAGCCATTATCTTAACACAGAT	480	
QY	681	AAAAAGTGTCAAAATCCCAAGACATTTTCCACTACTTGAGAGAATTTCTTAACCAAC	740	
QY	481	AAAAAGTGTCAAAATCCCAAGACATTTTCCACTACTTGAGAGAATTTCTTAACCAAC	540	
QY	741	CGTCAATTAATCTATTTGACAGATGATGGATGATGAAGCACTTCCAAACCTTGTCTTG	800	
QY	541	CGTCAATTAATCTATTTGACAGATGATGGATGATGAAGCACTTCCAAACCTTGTCTTG	600	
QY	801	AACAATTCGTGCTACTTTCATGATGGTGTGCTGCAAGCCGACAGATTTGGTATCTCT	860	
QY	601	AACAATTCGTGCTACTTTCATGATGGTGTGCTGCAAGCCGACAGATTTGGTATCTCT	660	
QY	861	CGTAAAGCTATGCTGAAGCATTTGCTATCTTGACAGTGTGTACAGTGAATACAGAGAT	920	
QY	661	CGTAAAGCTATGCTGAAGCATTTGCTATCTTGACAGTGTGTACAGTGAATACAGAGAT	720	
QY	921	CTAAGCACTGAATTAAGAATGCTACAAATGACAGCCCTTGAGACAGCTGCTGAAGATTA	980	
QY	721	CTAAGCACTGAATTAAGAATGCTACAAATGACAGCCCTTGAGACAGCTGCTGAAGATTA	780	
QY	981	GTTGATTAAGATAGCAGAGTAATGTTGAAGGTGCAAGGTTCAACAGCTATGCTTAAC	1040	
QY	781	GTTGATTAAGATAGCAGAGTAATGTTGAAGGTGCAAGGTTCAACAGCTATGCTTAAC	840	
QY	1041	CGATATGCACTGATTAATTCGCAATTAAGAACAACAACTTTCGACTTACCGTGAATAA	1100	
QY	841	CGATATGCACTGATTAATTCGCAATTAAGAACAACAACTTTCGACTTACCGTGAATAA	900	
QY	1101	CTACAAGAAAGCTTTGGCAAAATTAAGTGTGTGATGCTTTATCAAAAGTAGAGCTCA	1160	

Db	901	CTACAGAAACGGTTGGCAAAATTAGCTGGTGGTGTAGCTGTTATCCAAAGTAGGAGCTCCA	960
QY	1161	ACAGAGACAGCTTTAAAGAAATGAAACTTCGCAATGAGAGTGCCTPAATGCTACACGT	1220
Db	961	ACAGAGACAGCTTTAAAGAAATGAAACTTCGCAATGAGAGTGCCTPAATGCCACACGT	1020
QY	1221	GCAGCCGTTGAAAGAGATATGGTGGCTGGTGGTGGGAACAGCACTTATTAAGCTATTGAA	1280
Db	1021	GCAGCCGTTGAAAGAGATATGGTGGCTGGTGGTGGGAACAGCACTTATTAACGTCATGAA	1080
QY	1281	AAAGTAGCAGCTCTTGACCTTGAGGGGCGATGATGCTACTAGGAGCTTAACATTTGCTTCGT	1340
Db	1081	AAAGTAGCTGCTTGACCTTGAGGGGCGATGATGCTACTAGGAGCTTAACATTTGCTTCGT	1140
QY	1341	GCTCTAGAGAGACCTGTACGTCGCAAAATTGCTTTAAATGCTGGGTACGAAGGCTCCGTAGTT	1400
Db	1141	GCTCTAGAGAGACCTGTACGTCGCAAAATTGCTTTAAATGCTGGGTACGAAGGCTCCGTAGTT	1200
QY	1401	ATTGCAAGTTGAAAAAACAAGCCCTGCAGGAACAGAGATTTAATGCTGCAACAGGTGATGG	1460
Db	1201	ATTGCAAGTTGAAAAAACAAGCCCTGCAGGAACAGAGATTTAATGCTGCAACAGGTGATGG	1260
QY	1461	GTTGATATGATTTAAACAACGAGATCATTTAGCCCTGTCAAAAGTACACGATCAGCCCTCAA	1520
Db	1261	GTTGATATGATTTAAACAACGAGATCATTTAGCCCTGTCAAAAGTACACGATCAGCCCTCAA	1320
QY	1521	AATGACAGCTTCTGTAGCTAGTCTTATTTTGACAACAGAGCAGTTGTTGCTAATTAACCT	1580
Db	1321	AATGACAGCTTCTGTAGCTAGTCTTATTTTGACAACAGAGCAGTTGTTGCTAATTAACCT	1380
QY	1581	GAAACAGCTAGCCCAAGGCCACGACATGCAACAGAGTATGATTCAGAAATGATGGGTGG	1639
Db	1381	GAAACAGCTAGCCCAAGGCCACGACATGCAACAGAGTATGATTCAGAAATGATGGGTGG	1439

RESULT	2			
LOCUS	AF117741	2401 bp	DNA	BCF
DEFINITION	streptococcus pneumoniae strain CP1200 cochaperonin GroES (groES) and chaperonin GroEL (groEL) genes, complete cds.			
ACCESSION	AF117741			
VERSION	AF117741.1	GI:4566771		
KEYWORDS				
SOURCE	Streptococcus pneumoniae.			
ORGANISM	Streptococcus pneumoniae			
REFERENCE	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae.			
AUTHORS	Streptococcus.			
TITLE	1 (bases 1 to 2401)			
JOURNAL	Kim,S.N., Lee,J.Y., Kim,S.W., Choi,I.H. and Rhee,D.K.			
REFERENCE	groESL sequences in streptococcus pneumoniae			
AUTHORS	Unpublished			
TITLE	2 (bases 1 to 2401)			
JOURNAL	Kim,S.N., Kim,S.W., Choi,I.H. and Rhee,D.K.			
REFERENCE	Direct Submission			
AUTHORS	Submitted (04-JAN-1999) Pharmacy, Sungkyunkwan University, 300			
TITLE	Chunchun-dong, Suwon, Kyunggi-Do 440-746, Korea			
JOURNAL	location/Qualifiers			
FEATURES	1..2401			
SOURCE	/organism="Streptococcus pneumoniae"			
	/strain="CP1200"			
	/db_xref="taxon:1313"			
	122..603			
gene	/gene="groES"			
	122..127			
-35_signal	/gene="groES"			
	147..152			
-10_signal	/gene="groES"			
	262..288			
repeat_region	expression			
	/note="controlling inverted repeat of chaperonin			
	expression"			
	/rpt_family="CIRCE"			
	/rpt_type=inverted			

[illegible]

OY	494	CTCAGAAAGCTTGGAGGCTGGGGCAACAGTGGCTGGATTACCATGAGAAATCTCGAGG	55
Db	1098	CTCTGAAAGCAATGCAAAAAGTTGGCCAAAGACGGTGTCTATCCACCATGAAAGTCAACGGTGG	1155
OY	554	TATGAAAAACAGAACTGTAAGTGGTTGGAAGGCATGCATTTTGAACCGTGGTTTACCTGTCTCA	613
Db	1158	TATGAAAAACAGAGCTTGAAGTCGTAGAAAGAAATGCAAGTTTGAACCGTGGTTTACCTTTCACA	1211
OY	614	ATACATGGTCTACAGACAAATGAAAAATGGTTGCAGACCTTTGAAACCCATTATCTTAA	673
Db	1218	GTACATGGTGTCTATGATGCAAAAAAAATCGTGGCTGACCTTGAAAAATCCCTACATTTTGA	1277
OY	674	CACGATTAATAAAAGTGCACAAACATCCAGACATTTTGGCCACTCTCTGAGGAAGTTCTTAA	733
Db	1278	TACAGACAGAAAAATTTTCCATATCCAGAAATCTTGGCCACTTTTGGAAAGCATTTCTCA	1337
OY	734	AACCAACCGTCCATTACTCTATTTTCAGATGATGTGATGGTGAAGCACTTCCAAACCT	793
Db	1338	AAGCAATGTCTCACTCTTGATTTATTTGGATGATGTGATGGGAGAGCTTCCAAACCT	1397
OY	794	TGCTTTGACACAGATTTGCTGTACTTTCAATGTGGTTGCTGTCAAAAGCCACAGATTTGG	853
Db	1398	TGTTTTAAACAAGATTTGTGGAACCTTCAACGTAGACAGTCAAGGCAACCTGCTTTTGG	1457
OY	854	TGATCGTCGTAAAGCTATGCTTGAAGCANTGCTATCTTGACAGGGGTACAGTCATAC	913
Db	1458	TGACCGTGTCAAAACCCATGCTTGAAGATATGCGCACTTTACACGGGCGAACGTTATAC	1517
OY	914	AGAGATCTAGAGACTTGAATTAATAAGATGCTACAAATGACAGCCCTTGGACAGCGCTTAA	973
Db	1518	AGAAGACCTTGGTCTTGTATGATTAAGATGTGACAAATTTGAAGCTCTTGGTCAAGCAGCGG	1577
OY	974	GATTACAGTTGATTAAGATATGCACAGTAATTTGTTGAAGTTCAAGAGTTTCAAGACTAT	1033
Db	1578	AGTACCGTGTGACAAAGATATGCACAGTTATTTAGAGAGGTGACAGGAATCTGGAAGCAT	1637
OY	1034	TGCTAACCGTATTTGCACTGATTTAAATGCGAATTAAGAAACAAACCTTGCATTGACCG	1093
Db	1638	TTTCTACCGGTGTGGCTTATCAAGTCTCAAAATCGAACTTACACACTTCTGAAATTTTGACCG	1697
OY	1094	TGAATAAAGCTACAGAAAGTTTGGCGAAATTTAGCTGTGTGATAGCTTATCAAGTAGG	1153
Db	1698	TGAATAATTTGCAAGAAAGCTTGGCCAAATTTGTAGGTGTGTAGCGGTATTAAAGTTGG	1757
OY	1154	AGCTCCACACAGAGACAGCTTTAAAGAAATGAACCTTGCATTTGAGAGATGCTTAAATGC	1213
Db	1758	AGCGCGCACTGAAACATGAGTTTGAAGAAATGAACATCCCATTTGAAGATGCCCTCAACGC	1817
OY	1214	TACACGTGACCGCGTTGANAAGAAAGTATCGTGTGTGGTGGTGAACAGCACTATTATACGT	1273
Db	1818	TACTCGTGCACGCTGTGAAAGAAAGTATGTGTGAGAGTGTGGAAACAGCTCTTGGCCAAAGT	1877
OY	1274	TATTTGAAAAAGTAGACAGCTTGTGACGCTTGAAGCGCATGTGCTACAGCACTGAATGT	1333
Db	1878	GATTCCACGTGGAAGCTTACCTTGAATTTGACAGAGAAATGAAGCAACAGAGCAATATTTGT	1937
OY	1334	GCTTCTGTCTTAGAAGAGCCTTACGTCATAATTTGCTTAAATGCTGGGTACGAAGCTC	1393
Db	1938	TCTCCGTGCTTTGGAAGAACCCGTTGTCATAATTTGCTCACATGACAGATTTGAAGATC	1997
OY	1394	CGTGTATTATGACAAAGTTGAAAAACAGCCCTGACAGAACAGATTTAATGCTGCAACAGG	1453
Db	1998	TATTCGTATGATCGTTTGAAGAAATCTGAGCTTGGTATAGATTTTAAACGACGAAACGG	2057
OY	1454	TGAGTGGGTTGATATGATTAACACAGATCACTTGAACCTGTGCAAGTAAACAGATCGAC	1513
Db	2058	CGAGTGGGTTAACATGATTTGATTAAGGTTCACTTGTATCCAGTTTAAAGTGAATCGTTTACG	2117
OY	1514	GCTTCAAAATGACAGCTTCTGTAGCTAGTCTTATTTTGAACAAGAGAAGTGTGTTGCTTAA	1573
Db	2118	CTTACAAATATGACAGCATCTGTACCGAGCTTGAATTTTGAACAAGAGCAGTGTAGCCAA	2177
OY	1574	TAAACCTGAAACAGCTACGCCACGCGCCAGCAATGCCAGCAGGTATGATTCAGGAATGAT	1633

Db 2178 TAAACCCGAAACCACTAGCCCTCCAGCATGATCCAAAGCATGATGGCGGGATAT 2237

QY 1634 G 1634

Db 2238 G 2238

RESULT		3	
LIGROESSEL			
LOCUS	LIGROESSEL	2267 bp	DNA
DEFINITION	L. lactis groES and groEL genes.		BCF
ACCESSION	X71132		
VERSION	X71132.1 GI:287869		
KEYWORDS	groEL gene; groES gene; heat shock protein.		
SOURCE	Lactococcus lactis.		
ORGANISM	Lactococcus lactis		

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 2267)
Kim, S.G. and Batt, C.A.
Cloning and sequencing of the *Lactococcus lactis* subsp. *lactis*. .
Gene in press
2 (bases 1 to 2267)
Batt, C.A.
Direct Submission
Submitted (04-MAR-1993) C.A. Batt, Cornell University, 107 Stocking
Hall, Ithaca, NY 14853, USA

```

FEATURES
  source      Location/Qualifiers
              1..2267
              /organism="Lactococcus lactis"
              /strain="ssp. lactis"
              /isolate="CC9"
              /db_xref="taxon:1358"
misc.feature 194..220

```

```

gene      246..530
          /gene="groES"
CDS       246..530
          /gene="groES"
          /codon_start=1
          /transl_table=11
          /protein_id="CAA50445.1"
          /db_xref="GI:287870"
          /db_xref="SWISS-PROT:P37283"
          /translation="MLKPLENRVLRKEREKESMGGIVLTSAQEKPTAEVAVGE
          GKTNHGGTISPLVKKGVDFVIFKKPAGTIVKMGGEFLIKDSDLAIYE"
gene      618..2246
          /gene="groEL"
CDS       618..2246
          /gene="groEL"

```

```

/gene _glocr-
/codon_start=1
/transl_table=11
/protein_id="CAA50446.1"
/db_xref="GI:287871"
/db_xref="SWISS-PROT:P37282"
/translation="MSKDKIFSSDARTAMRGDILADIYKTTFLGPKGRVILEKSYG
SPILINDGVITAKEIELEIEVHFENMGAKLVEASKNDINDAGDGTNATVATVIGEG
LNKVTAGANPVGIRIGRIELAEFTAVASIKEMALPVHDKSAIAOVAVVSSSEKVGRI
SDMEEVGSNDQVITIEESKQMOTELDVBEWEDDRCLYQSVMYKGVNFAEINDPYI
LITDKISNIOETLPELQILKTNRPDLVADVDEALPTVLNLIKGVNFAVAKA
PGSGDRRRAKOLEHLALTGTGTVITEELGDLNKATLEALGOAKATVDHDTIVGA
GADALSDRAILIKAOIEKRTSDPREKIOERAKTAGVAAVKAAVADNRETKAKML
LIEDALNATRAAEVSGIVSGCAITVNAITLADKLSEEDIGCGINTVRALEPFRQ
ILNACVYEGSVIIDKLRESEVGTGFMAATGQVNAITLSEGLVDPAKVTASLONAA5VA
GALLTEAVVANNPEEPAAMPMDMSKMGMI"

```

Query Match	57.1%	Score 948.6;	DB 2;	Length 2267;
Best Local Similarity	74.9%	Pred. No. 1.3e169;		
Matches 118;	Conservative	0;	Mismatches 399;	Indels 0;
			Gaps	0;

OY	11	TCATATGGCAAAAGAAATTCAAATTTTCAGAGATGGCGCGTGGCAATGGTCCCGAGT	70
Db	614	TAAATATGCAAAAGATATTTAAATTTTCAGTATGCTACAAACAGCATGATGCTGGGAT	673
OY	71	TGATATGTTAGCAGATACCGTCAACAGTAACGCTTGGCTCTAAAGGCGCAATGTGTTCT	130
Db	674	TGATATTTCTGCTGATACAGTAAAAAACACCCCTGGACCAAAAGTCGCAATGTGTTCT	733
OY	131	TGAAAAAGCTTTGGTCTCCCTTAATTCTATATGACGGGGTAACCTTTGCTAAAGAT	190
Db	734	TGAAAAATCATATGATACCTTTAATTAATACAAACGATGGTGTAAACGATTTGCCAAAGAT	793
OY	191	CGAATTTAGAAATCATTTTGA AAAACATGGAGCAAAATTTGCTGTGAAGTGCTGTAA	250
Db	794	TGAATCTGGAATACATTTTGGAAAAATATGGAGCTAAACTGTGTTACAGAAATGCTTCAA	853
OY	251	AACCAATGATATTTCTGTGTATGGAGCAGACTACTGCACAGTTTGTACACAACCATTTGT	310
Db	854	AACCAATGATATTCGACGTACGCGTACAAACACAGCGACAGTTTGTACACAAGCTATCT	913
OY	311	TCATGAAGACTAA AAAATGTGACAGCAGGTGTATTCCAATTTGGATTCGCGAGGAT	370
Db	914	TCGTGAAGTTTAAAAATGTTTACCGCAGGTGCAATTCAGTTGGTATTTCCCGCAGGAT	973
OY	371	TGAAAACAGCAACAGCACAGCTGTTGAAGCCTTGAAGCCATTTGCTCAACCTGTATCG	430
Db	974	TGAATCTGCTCTGAACACACTGTTGGGTCATTAATTAAGATGGCAATTTCTGTCCAGA	1033
OY	431	CAAGAAAGCTTTGCTCAGTCCGCTGCAGATCATACAGCTCTGAAAAAGTTGGAGATA	490
Db	1034	TAAATTCACCAATTTGCGCAAGTACCTACCGTTTATACCTAGTAAAAAGTGGGTGAATA	1093
OY	491	TATCTCAGAAAGCTATGGAGGCTGTGGGCAACGATGGTGTATTAACATGCAAGATTCG	550
Db	1094	TATTTCTGATGCATATGGAAGTGTAAGTTCTGACGGAGTTATCACCATTTGAAGATCAA	1153
OY	551	AGTATGGAACAGACTTGAAGTGGTTGAAGSCATGCAATTTGACCGTGGTATCTGTC	610
Db	1154	AGGATGCAAACTGCAACTCATGTGTTGGAAGAAATGCAATTTGACCGGTATTTAG	1213
OY	611	TCATATACCTGCTCACAACATGAAAAAATGGTGTGCAGACCTTGA AAAACCCATTATCT	670
Db	1214	TCATATATATGTTCTTATATACAAAAAATGGTGTGATTAAGATATCTTATATTTCT	1273
OY	671	AATCAGGATAAAAAGTGTCAAAACATCCAGACATTTTGCCACTACTTGAAGAGTTCT	730
Db	1274	TATTAACGACAAAAAATCTCAAAACATTCGAAGAAATTTTACCGCTCTGTGAACAAATCTT	1333
OY	731	TAAAAACCAACGCTCATTTACTCATTTATGCAATGATATGGATGGTGAAGCACTTCCAC	790
Db	1334	GAAAAACAAATGCTCACTCTTTATTTAGTGTGATGTTGATGGAAGAAAGCATTTGCCAAC	1393
OY	791	CGTTGCTTGAACAAGATTTGCTGTACTTTCAATGTGGTGTGTCGCAAGCCGACGAT	850
Db	1394	GCTTGTCTTATTAAGTTAAAGTGTCTTCATGTTAGTGTGTAAGAACCAACGAT	1453
OY	851	TGTGATGCTGTAAAGCTATGCTTGAGACATTTGCTATCTTGACAGGTGTACAGTAT	910
Db	1454	TGTGATGCTGTAAAGCTCAATTTGACACTTTGGCTATCTTGACAGGTGTACAGTAT	1513
OY	911	TACAGAGATATAGACTTGAATTTAAAGATGCTACAAATGACAGCGCTTGAACAGCGC	970
Db	1514	TACTGAAAGAACTTGGTCTTATCTTTAAGATGCTACTCTTGAAAGCTTTTGAAGCAAGC	1573
OY	971	TAAGATTTACAGTTCAATAAAGATAGCACAGTAAATTTGTTGAAGGTTTCAAGAAAGTTGAGAAGC	1030
Db	1574	TAAAGCAACTGTGATTAAGACCAACACAACAATTTGTTGAAGGTGCGTGTGCTGATGC	1633
OY	1031	TATTTGCTAACCGTATTTGCATGATTAATTCGCAATTAAGAAACAACTTCTGACTTTGA	1090
Db	1634	TATTTGCTGATGCTACTTTATTTAAAGGCGCAAAATTTGAAAAACAACTTCTGATTTTGA	1693

QY	1091	CCGGAAAAACTACAAAGACGTTGGCGAAATTAAGTGGTGGTGAAGCTGTATCAACAT	1150
Dp	1694	TCGCGAAAAAATTACAAAGAACGCTGTGCAAAATTAGCTGGTGTTCGACATGTTAAAGT	1753
QY	1151	AGGAGCTCCACAGAGACAGCTTTAAAGAAATGAAACTTCGATTGGAGATGCTTAA	1210
Dp	1754	TGCGTGCACGACATGAAACACAGAGCTCAAAACCAATGAACACTTGATTGAAGATGCATTA	1813
QY	1211	TGTCACACGACGACCGGTTGTAAGAGATTCGTGGTGGTGGAGACAGACTATTAC	1270
Dp	1814	TGCACACGTCGACCGCTTGAGAACAGTATTGTTTCTGGTGGTGAACGACACTTGTAA	1873
QY	1271	GGTTATTGAAAAAGATGACAGCTCTTAGCTTGAGGCGATGATGCTACTGACGTAACAT	1330
Dp	1874	TGCTATTGCTGCTTTGGATTAACTTTCACAGAGAGAGATATTCAAACAGTATTAAAT	1933
QY	1331	TGTCCTTCGAGCTCTGAGAAAGCCTGACGTAATGCTTAATGCTGGGTGAGAAAG	1390
Dp	1934	TGTTCCCGCTGCTCTTGAAAGAACACAGTGTGCTCAATGCTGCCAATGAGGATATGAAAG	1993
QY	1391	CTCCGATGATTATGACAAAGTTGAAAAACAGCCCTGCAGGAAACGATTTAATGCTGCAC	1450
Dp	1994	TTTCAGTATATCATGATTAACTTGCTTCACAGAGAGATGATGATTCATATGCTGCAC	2053
QY	1451	AGGAGATGGGTTGATATGATTAAACAGGAATCATGACCTGTCCAAAGTAACACGATC	1510
Dp	2054	TGGCAATGGGTTAAATATGATTGAAAGAGAAATGTGACCCCTGCAAAAGTAACCTGTC	2113
QY	1511	AGGCTTCAAAATGACAGCTCTGACCTAGTCTATTTTGAACAGAGAGATGTTGGT	1570
Dp	2114	AGCTTTCAAAAGCGGCTTCTGTGCTGGTTTAATTTGACTACTGAAGCAGTATGTTGC	2173
QY	1571	TAATTAACCTGAAACGACTACGCGAC	1597
Dp	2174	TAATTAACCGCAACCACTGCTCCAC	2200

FEATURES	source
LOCUS	AB028452
DEFINITION	AB028452 2834 bp DNA BCT 07-DEC-1999
ACCESSION	AB028452
VERSION	AB028452.1 GI:6526959
KEYWORDS	Cpn60; Cpn10.
SOURCE	Bacillus sp. MS (strain:MS) DNA.
ORGANISM	Bacillus sp. MS (strain:MS) DNA.
REFERENCE	Bacteria; Firmicutes; Bacillus/Clostridium group;
AUTHORS	Bacillus/Staphylococcus group; Bacillus.
TITLE	1 (bases 1 to 2834)
JOURNAL	Kondo, A. and Yohda, M.
REFERENCE	Cpn60/10 from Bacillus strain MS
AUTHORS	Published Only in Database (1999) In press
TITLE	2 (bases 1 to 2834)
JOURNAL	Kondo, A. and Yohda, M.
REFERENCE	Direct Submission
AUTHORS	Submitted (03-JUN-1999) to the DDBJ/EMBL/Genbank databases. Akihiko
TITLE	Kondo, Kobe University, Department of Chemical Science and
JOURNAL	Engineering, 1-1 Rokkodaicho Nada, Kobe, Hyogo 657-8501, Japan
REFERENCE	(E-mail:kondocx.kode-u.ac.jp, URL:http://www.kode-u.ac.jp/,
AUTHORS	Tel:078-803-6196, Fax:078-803-6206)
TITLE	Location/Qualifiers
JOURNAL	1..2834
REFERENCE	/organism="Bacillus sp. MS"
AUTHORS	/strain="MS"
TITLE	/db_xref="taxon:96470"
JOURNAL	534..818
REFERENCE	/gene="Cpn10"
AUTHORS	534..818
TITLE	/gene="Cpn10"
JOURNAL	/codon_start=1
REFERENCE	/transl_table=11
AUTHORS	/product="Cpn10"
TITLE	/protein_id="BAA88109.1"

```
/db_xref="GI:6526960"
/translation="MLKPLGDRVIVIEIETEEKTASGIVLPDTAKERPOEGRVAVAGK
GRVLSGEVVEVEVDRIRIFSKVAGTEVKYDKKEYLILRESDLIAVIG"
gene
901..2520
/contig="Cpn60"
/contig_start=1
/contig_end=11
/product="Cpn60"
/protein_id="BAA8110.1"
/db_xref="GI:6526961"
/translation="MAKEIKFSEBRNRAMLRGVDKLADAVKYTLGPKRNVLEKREG
SPILINDGVYIAKEIELEDPEFENMGAKLVAEASKINDVAGDGTATVLAOAMIREG
LNRTVAGAMPNGIRKIEKAVAVAEELAIKIPKESISIOVAISADEVEQOLI
AAEMERGVNDGVYTLSEKSPTELDVVEGMOFDRGVSPYMTTDEKMEVLENYI
LITDKISNIDILPLEQVOOGKPLLIADIEBEALATLVNKLGRFPVAVAKA
PGSDPRKAMIEDIALITGEVYSEPELSEIKSATIASGRASKVYVTEENTIVGA
GDSERKARINIRAOLETTSEFDEKLOERLAKLAVGRVYIKVGAATELEKRL
RIEDALNSTRVAVEGIVAGGFTALNNVKNVALEDEDAIVKIVLRAIEEPVRO
IAONAGLEGSVIERLTERKPGIGENAAIGEVNDMEAGIVDPTRKVSRLONAAASA
AMFLTEEAVVAAPKEENKCNAGMPDMGMM"
BASE COUNT 954 a 525 c 668 g 687 t
ORIGIN
Query Match 50.6%; Score 840.4; DB 1; Length 2834;
Best Local Similarity 71.0%; Pred. No. 3.4e-149;
Matches 1114; Conservative 0; Mismatches 456; Indels 0; Gaps 0;
QY 14 TATGGCAAAAGAAATCAATTTTCAGCAGATGCGCGTGTCCATGTTGGCGGAGTTGA 73
DB 900 TATGGCAAAAGAAATTAATTAATTCAGCAGAGAGCTGTCGTCAGTGTACGTGTTGCGTGA 959
QY 74 TATGTTAGCAGATACCGTCAAGTGGTGGTCTTAAAGGGGCGCAATGTTGTTCTTGA 133
DB 960 CAATTTAGCTGATGCACTAAAGTAATACGTTAGGTCCTAAAGGCGGCTTACGTCGATTAGA 1019
QY 134 AAAAGCTTTTGTTCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 193
DB 1020 GAAATAATTCGTTCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1079
QY 194 ATTAGAAGATCATTTTGAAGACATGGAGCAAAATTTGTTGAGTGGCTCTTAAC 253
DB 1080 ATTAGAAGATCATTTTGAAGACATGGAGCAAAATTTGTTGAGTGGCTCTTAAC 1139
QY 254 CAATGATTTGCTGTGATGAGGAGCACTACTGCAACAGTTTGAACACCAATTTGTTCA 313
DB 1140 AAACGATGTTGCTGTGAGGAGCGTACACACGCAACAGTTTGAACACCAATTTGTTCA 1199
QY 314 TGAAGACTAAATAATGTGACAGCAGGTGTAATCCAAATTTGTTGCTGAGGCAATTTGA 373
DB 1200 CGAAGACTTGAATAATGTGACAGCAGGTGTAATCCAAATTTGTTGCTGAGGCAATTTGA 1259
QY 374 AAAGCAACGCAACAGCTGTTGAAGCTTGAAGCATTGTCATCCATCTATCTGGCAA 433
DB 1260 AAAAGCGGTTGCTGTGCACTAGAGATTAAGCAATTTCAACCAATTTCAAGGAAA 1319
QY 434 GGAAGCTATTGCTCAGGTGCTGCTGATCATCATGCTGGAAGAAATTTGAGAGATTTAT 493
DB 1320 AGAATGATTTGCTCAGGTGCTGCTGATCATCATGCTGGAAGAAATTTGAGAGATTTAT 1379
QY 494 CTCAGAACTATGAGCGCTGTGGCAACGATGTTGATTAACATCGAAGATTTCTGAGG 553
DB 1380 CCCAGAGCAATGAGCGCTGTGGCAACGATGTTGATTAACATCGAAGATTTCTGAGG 1439
QY 554 TATGAAACAGACTTGAAGCTGTTGAAGCATTGACCGTGTACCTGCTCA 613
DB 1440 CTTCACACAGCAATTTGATGTTGTAAGGTATGCAATTTGACCGTGTATGATCTCC 1499
QY 614 ATACATGTCACAGCAATTTGATGTTGTAAGGTATGCAATTTGACCGTGTATGATCTCC 673
DB 1500 ATACATGTCACAGCAATTTGATGTTGTAAGGTATGCAATTTGACCGTGTATGATCTCC 1559
```

```
QY 674 CACGATMAAAAGTGTCAAAACATCCAAAGCATTTTGCATCTTGAAGAGTTCTTAA 733
DB 1560 CACAGACAAAATAATCTCTAACAATCCAAAGCATCTTGCCTATCTTGAACAAGTTGTCA 1619
QY 734 AACCAACGCTCATCTACTCTTATATGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 793
DB 1620 AACCAACGCTCATCTACTCTTATATGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1679
QY 794 TGCTTGAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 853
DB 1680 AGTAGTAAACAACTTGCCTGCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 1739
QY 854 TGATGCTGTAAGCATATGCTTGAAGCATTTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 913
DB 1740 TGATGCTGTAAGCATATGCTTGAAGCATTTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1799
QY 914 AGAGATCTAGCACTTGAATTAAGATGCTACAAATGACAGCCCTTGGACAGCTGTAA 973
DB 1800 CGAAGAGCTAGGAGCTGGAATTAATCTGCAACATCTGATCTGATCTGATCTGATCTGATCTGATCTG 1859
QY 974 GATTACAGTGTGATTAAGATGACAGATTAATGTTGAAGGTTGCAAGAGTTGACAGAGCTAT 1033
DB 1860 AGTAGTAAACAACTTGCCTGCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 1919
QY 1034 TGCTAACGCTATTTGCACTGATTAATGCAATTAGAAACAAACATTTGATCTGATCTGATCTGATCTG 1093
DB 1920 TAAAGCTGCTATACCAACCAATTTGCTGCTCAATTAGAAACAAACATTTGATCTGATCTGATCTGAT 1979
QY 1094 TGAATACTCAAGAAAGCTTTGGCGAAATTTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1153
DB 1980 TGAATACTCAAGAAAGCTTTGAGAAATTTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2039
QY 1154 AGCTCCACAGAGACAGCTTTAAAGAAATGAAACTTCCGATTTGAGAGATCTGTAATGC 1213
DB 2040 TGAAGCTGACGAAACAGAAATTTGAAGAGCAATTTAGCTGATGATGATGATGATGATGATGATGATG 2099
QY 1214 TACAGCTGACGCGTGTGAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1273
DB 2100 TACTGCGCGCTGTGGAAGAGATGCTGAGCGCGGCTGCTGAGCAATTAATGAACGT 2159
QY 1274 TATTGAAAGATGACAGCTTTGAGCTTGAAGGCGATGATGATGATGATGATGATGATGATGATGATG 1333
DB 2160 ATACAAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2219
QY 1334 GCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1393
DB 2220 TCTTCGCGCAATTTGAAGAGCACTTCCCAATTCGCCCAAAAGCGCTGCTGCTGCTGCTGCTGCTG 2279
QY 1394 GCTGATTTGACAGCTTGAAGAAACAGCCCTGAGAGCAAGATTTAATGCTGCAACAG 1453
DB 2280 TGTCAATTTGTAAGAGCTTTAAAGCAAGAAACCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTG 2339
QY 1454 TGAATGCTGATGATGATTTAAAGCAAGAAATGATGATGATGATGATGATGATGATGATGATGATG 1513
DB 2340 TGAATGCTGATGATGATTTAAAGCAAGAAATGATGATGATGATGATGATGATGATGATGATGATG 2399
QY 1514 GCTTCAAAATGACAGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1573
DB 2400 ACTTCAAAAGCAAGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2659
QY 1574 TAAACCTGAA 1583
DB 2460 CAAACCAAGAA 2469
RESULT 5
BACSGROESL 2309 bp DNA BCT 22-FEB-2000
LOCUS Bacillus stearothermophilus GroES (groES) and GroEL (groEL) genes,
DEFINITION complete cds.
ACCESSION L10132
VERSION L10132.1 GI:289298
```



```

LITDKITINIOELIPYLEQVYQGGKPIILLIAEDVEGALATLVVNNKLKGFENAVAKA
PGFDRRRKAMLEIAVLTLGEVITEDIGLIDKSTQIOLQGRASKVVVTKENTTIVEGA
GETDJKISARVQITRAQVEETSEFDRKIQERLAKLAGVAVIKVATETELERKL
RIEDALNSTRAVEEGIVSGGTALVNVNVAALVNAVEGDAOTGINTIVRLALEPIRO
IAHNAGLEGSIYERLKEEIGVFNATGEMVAMIKGIVDPFKYTRSAIQNNAASVA
AMPLTTEAVVADKREENGCGAGMPDMGCMGCMGM"
RBS      3506..3514
/gene="ydlM"
/gene=".3900
3506..3900
/gene="ydlM"
3520..3900
/gene="ydlM"
/note="function unknown"
/codon_start=1
/transl_table=1
/protein_id="BAA22748.1"
/db_xref="GI:2521995"
/translacion="MISDNFVSETELENNVOYLLPGAKAVIKRGYSRAHKYTTAKSEV
EKMVAAGLSGSEIAILLIRKHWIGASIPAGGIVIDIDSKOGELVGLLDAQNPFCH
CIRTLMGSLFLRITNMGRKRLNK"
RBS      5056..5061
/gene="ydlN"
5056..5282
/gene="ydlN"
5067..5282
/gene="ydlN"
/note="function unknown"
/codon_start=1
/transl_table=1
/protein_id="BAA22749.1"
/db_xref="GI:2521996"
/translacion="MIKFSVTLGIMRCSLTHTTTKNTNALKRMITYPKRPSFFHEEK
VLXKLKFKICIKIMIKIRSCMGYFL"
RBS      5301..5307
/gene="ydlO"
5301..6595
/gene="ydlO"
5312..6595
/gene="ydlO"
/note="similar to modification methylase"
/codon_start=1
/transl_table=1
/protein_id="BAA22750.1"
/db_xref="GI:2521998"
/translacion="MTNFIENENKOLSLATEDENIENFYIDGTDVIRKTIIRSGSGVT
SRVPLSTODLENNKNTHELYDESMILBKNRPNTLETSEINADIPEGCGGLSGAYE
ACRALGINRPFACDLNEALSYEKNRSPDLSNRSIKHINDELGAFLIVEORNI
KDRVVKRIDFTLAGPCQGHSDLNHNRKDPRLNALMRVSRVIELFQPSVVLVENQI
I1HDKSGSFKEFNHLKTOGYRDEIVLNAEKLGVSGARRRYEFAKSTPVSSLNQIN
EFTJNSRPIYSMAISDLVENVGDDIFNTASEHLENKRRIEYLFENNLFELPNSERPD
CHRLKPSYSKSYGRMYWDRPAPLTITRGFGSTGGRVHSLKRTIIPHEARLRQFFP
DFPNFDLRRROYDVIIGNAVPSKLSYLLALHQLR"
RBS      6604..7786
/gene="ydlP"
6604..6611
/gene="ydlP"
6617..7786
/gene="ydlP"
/note="probable modification methlase"
/codon_start=1
/transl_table=1
/protein_id="BAA22751.1"
/db_xref="GI:2521998"
/translacion="MKVVSLEFSGIGIEGLHOSGHTTEIFCEYDPLAKAVLSKNFPG
VKTEDINEIRELPSCDLYAAGPCQDLSQAGKEGIDGSRSGLVKFLFLEIKHEHA
NRPPWILIEVNPMLRLNKGAMSYLTSYLGTYWARYTVADRCFGLQRRRIKRL
LASLFEDPKDVIYSODHSEPDLDGKPSVDHNSYVYGTGELRGVGMAREAVPIC
GSSVGIASPAVMSPYEDIYGTINIRDAERLOGEPEDMTNITTEGDKIDEGARMLY
GNAVSVFVSKWIGENLSOPKGSISDEGELVTKTPWSAAGYGDKKYKVPVSKWVAT
EOTIAISEPLNHPKLPLSARLNLNGLGRAACTVANSDEPINSLERCKDROLQV"
terminator
RBS      8255..8262
/note="stem loop"

```

```

/gene="ydlO"
8255..8460
/gene="ydlO"
8272..8460
/gene="ydlO"
/note="function unknown"
/codon_start=1
/transl_table=1
/protein_id="BAA22752.1"
/db_xref="GI:2521999"
/translacion="MKSFELIVSTYLNCKDIRLNSILVSKRVANOIALAFVPEEIM
FSCUVAINISGNEIKSR"
RBS      9697..10653
/gene="ydlR"
9697..9704
/gene="ydlR"
9712..10653
/gene="ydlR"
/note="function unknown"
/codon_start=1
/transl_table=1
/protein_id="BAA22753.1"
/db_xref="GI:2522000"
/translacion="MPEIKRLEDAVEYLLGNYPAGVSTSYSKYNEIRKIYSEAPLIG
ENEYVVTGTSRIRISLIDTRFACQNRKHTMAVAFTSIPVDSSELELIIRTRNSA
ITTSKPCNGEERCTIFDGLFLVPGEGEVYPLAFVPENDSELEKKAPELLPGTE
LKEVPRANOSPACQENNKSAKNEDESKSVYFELIDEDGSIYFVEDKRTYRIGM
IWTASHNNGSSAITRLEVELEVENVLVCKIKHKYNEVPDKNLSLKFVNIQDLISFL
DLHPNVQSGSEFVSDIVDENATSSDDLPEDENN"
RBS      10668..10674
/gene="ydlS"
join(10668..10674,10685..11716)
/gene="ydlS"
10685..11716
/gene="ydlS"
/note="weakly similar to 5-methylcytosine-specific
restriction enzyme McrB of E. Coli"
/codon_start=1
/transl_table=1
/protein_id="BAA22754.1"
/db_xref="GI:2522001"
/translacion="MEISQTSDDLISLEKKKGTLPPKFSVLSIRPNRIIYAGPTGK
SNYLEREVGKIFGDNPVYFTRVFPFGYTGOFIGAVKPDVYKLSGEEIESSNR
DKMENPEPMIDYOPGPGPTDVLKALKNRYTNLIIETIRNANASVPGDIPQLD
RNKGSDEYVTECPDILMNTLARGIDEMAIKLPDSNFFIMATNNADQVLPDITAK
RMSFEYLELEKTRKAVDSWKLSTRYGHNKVIMWNPFDIIRLRLKQVPEKLLGP
FELKESELMOQVFNKRLLYLKEDVEKHNPTIDFLNASTFSELEIERYDGSNDNIFPD
IDSSFEVSD"
RBS      11705..11712
/gene="ydlA"
11705..11712
/gene="ydlA"
11719..13116
/gene="ydlA"
11719..13116
/gene="ydlA"
/note="function unknown"
/codon_start=1
/transl_table=1
/protein_id="BAA22755.1"
/db_xref="GI:2522002"
/translacion="MDKSSKFFPEDOKYNNERIVRIGNALALSKGILYEDSSGCL
IFNVGYISGRNVITFLPKYCNHRHDEHSKRTLEFNKLITFKYSGLNLSRSDYFV
SELDSDSDVDFMNLADYLLNDPSLNGYQOKRTEYIDEGEIIIMSKTVNBITPVSFG
VPPYFSTYNEVVQODEYHLIVKIKHMAKLSYFNDFGVILGFTGLEPFRSCDGMILDY
ADFGYSVINKELIVWTYVDYDYLKALKTAIDRENOFSKRPGLSLGTGYPHRWME
VCKTQFHVHVAWYKISRPMMINPTDILEVKKKKTLEPDITKAPEYSKRTYPLIDAK
YXNINPFGKLECNPGVEDTETKOLYVLALEKLSRGTTRKNAFLFPSSNRTNFKFVG
SVDPDPLDIAVTLVITSABOVYMLYENKTFSTDLKFPVSEINKSKRKHVSITSTL
YGNMFLTKRLSDKN"
Query Match      48.6%; Score 807.4; DB 1; Length 17516;
Best Local Similarity 68.0%; Pred. No. 4,1e-143;

```

Matches 1126; Conservative 0; Mismatches 531; Indels 0; Gaps 0;

```

QY 1 GAATGCGCTTCATATGGCAAAAGAAATCAATTTTCAGCAGATGCGCGTGCATGG 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 308 GAGGCTTGTAAACATGGCAAAAGAAATTAAGTTTAGTGAAGAGCGCGCGGCATGC 367
QY 61 TGGCGGAGTGTATATGTTAGCAGATACGCTCAAGAGTGGCTTGGCTTAAAGGGCGCA 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 368 TTGCGCGGTGTCAGTGTGATGCTGTAAAGTAACTTTAGAGCAACCAAGGAGCGCA 427
QY 121 ATGTTTCTTGAAGAGCTTTGGTTCTCCCTTAATTAATTAATGAGCGGGGTACCATG 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 428 ACCTGCTTACAGAAAAATTTGGTTCTCCCTTAATTAATTAATGAGCGGTACCATG 487
QY 181 CTAAAGAGATCGAATTAAGAGATCATTTTGAAGATGGAGCAAAATTTGGTGTGAAG 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 488 CTAAAGAAATCGAGCTAGAGAGCGGTTGAAGAAATGGGTGTGAAGCTTGTGTGAAG 547
QY 241 TGGCTTCTAAACCAATGATATTGCTGTGATGGAGCGACTACTGCAACAGTTTGCACAC 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 548 TAGCCGAGAAACAAACGAGCTGCGCGGTACACACACACACAGCTTGTGCGC 607
QY 301 AAGCCATTTGTCAGAGGCTTAAAGATGTGACAGAGGTGCTAATCCAAATGGTATCG 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 608 AAGCAATGATCCGTGAAGGCTTAAAGCTTAACAGAGCGGCTTAACCTGTAGCGCTGC 667
QY 361 GTGAGGCAATTAAGACAGCAACAGCTTTGAAGGCTTGAAGCCATTGCTCAAC 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 668 GTAAAGGATGAGCAAGCTGTAGCGTTGCGATCAAACTTAAAGAAATTTCTAAGC 727
QY 421 CTGTATCTGGCAAGAGCTATTGCTCAGGTGCTGAGTATCATACGCTTGAAGAAAG 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 728 CATTCAGAGCAAGAGCTATGCTCAGGTGCTGAGTCTGCTGTGATGAGAGAG 787
QY 481 TTGGAGATATATCTAGAGCTATGAGCGGTGGGCAACGATGGTGTATACCATG 540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 788 TGGGAAGCCTTATCGCTGAAGCAATGAGCGGCTAGGAAACGCGCTTATACATATG 847
QY 541 AAGAAATCGAGGTATGAAACAGAACTTGAAGTGTGAAGGATGCAATTTGACCGTG 600
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 848 AAGAGCTTAAAGGCTTCAACACTGAGCTTGAAGTGTGAAGTATGCAATTCGACCGG 907
QY 601 GTTACTGTCTCAATGATGCTCAGACAGACATGAAAAATGTTGCAAGCTTGAAGAC 660
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 908 GATATGCGTCTCTTACATGATGATGATGATGATGATGATGATGATGATGATGATG 967
QY 661 CATTTATCTTAATCAGAGTAAAGAAAGTCAAAACATCAAGACATTTTGGCCACTTGG 720
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 968 CTTCATCTTAACTACAGACAAAAAAATCAACAACATTCAGAAATCTTCTGCTGCTTG 1027
QY 721 AAGAACTTCTAAACCAACCGTCCATTAATTAATGCAAGATGATGATGATGATGATG 780
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1028 AACAGGTTGTGACGAAGCAAAACCTTGTCTGTATGCTGATGATGATGATGATGATG 1087
QY 781 CACTTCCAAACCTGTGCTTGAACAGATTCGTGCTTCAATGCTGCTGCTGCTGAAG 840
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1088 CACTTCTCACTGTGTTGGAACAACTTCGCGGCACTTCAACGAGAGGCGCTTAAAG 1147
QY 841 CCGCAGAGTTTGGTATCGTATGATGATGATGATGATGATGATGATGATGATGATG 900
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1148 CTCTGCTTTCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1207
QY 901 GTACAGTATTAACAGAGATCTAGACCTGTAATTAAGATGCTACATGACAGCCCTTG 960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1208 GAGAAGTCTATCAGAGAGATCTTGGCTTGAACCTTCAATCTCAATGCTCAATG 1267
QY 961 GACAGGCTCTTAAGATTAAGTTGATGAAGATGACAGCAATGTTGAAGTTCAAGAA 1020
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1268 GAGCGGCTTCAAGTTGCTGATTAAGAAACCAACCAATCGTTGAAGGCGCTGAGG 1327
QY 1021 GTTCAAGAGCTATGCTTAACCGTATGCACTGATTAATTCGCAATTAAGAAACCACTT 1080
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1328 AAACAGACAAATTTCTGCGCGGTGCTGCTCAATCCCGCTCAAGTGAAGAAACCACTT 1387

```

```

QY 1081 CTGACTTTGACCGTGAAAACTACAGAACGTTTGGCGAATTAGCTGTGTAGCTG 1140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1388 CTGAATTCGACAGAGAAAAATTAACAGAGCGCTTGTGTAACCTTGTGCGGCTGAGCTG 1447
QY 1141 TTATCAAGTAGAGCTCCACAGACAGACACTTTAAAAAGAAATGAATCTTGCATTGAGG 1200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1448 TCATCAAGTCGCGGCTGCGACAGTGAAGAGCCGTAAACCTTCGATGGAAG 1507
QY 1201 ATGCTCTAAATGCTACACGTCAGCCGTTGAAGAGATGCTGCTGCTGCTGCTGGAACAG 1260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1508 AGCGCTTGAACCTAACCTCGCGCAGCTGTGTAAGAGCATGATCCGCTGCTGCTGAG 1567
QY 1261 CACTTATTAAGGTTATTAAGAAAGTAGACAGCTCTTGAAGCTTGAAGGCGATGCTACTG 1320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1568 CGCTTGAACGATATTAACAAAGTCGCTGACGTTGAAGCTGAAGGAGATGCTCAAAACAG 1627
QY 1321 GAGCTAACATTTGCTTCTGCTGCTGCTGCTGAGAGAGCTGACTCAATTTGCTTAAATGCTG 1380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1628 GTATCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1687
QY 1381 GGTACGAGAGGCTCCGTAGTATTAAGCAAGTTGAAGAAACAGCCCTGCAAGAGAGATTTA 1440
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1688 GCTTTAAGAGATCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1747
QY 1441 ATGCTGCAACAGGCTGAGTGGTGTATGATTAATTAAGAGAGATCAATGACCTGTCAAG 1500
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1748 ACCTGCAACCTGCGATGAGTGGTAAACATGATGCAAAAGATATGCTGACCAACAAAG 1807
QY 1501 TAACAGATCAGCGCTTCAAAATGCAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1808 TTACAGCTCAGCTGCTTCAAAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1867
QY 1561 CAGTGTGCTTAATTAACCTGACAGCTACGCGACGCGCAGCATGCGCAGCGATGAG 1620
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1868 CTGTTGCTGCTGCAACCGCAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1927
QY 1621 ATTCAGAAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1657
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1928 GCATGGCGGATGAGTGAATGATGATTAAGAGGA 1964

```

RESULT 9
BSUB0004
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

BSUB0004 213190 bp DNA BCT 26-NOV-1997
Bacillus subtilis complete genome (section 4 of 21): from 600701 to 813890.
299107 AL009126
299107.1 GI:2632866

Bacillus subtilis.
Bacteria: Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Bacillus.
1 (bases 1 to 213190)

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G.,
Azevedo, V., Bertero, M.G., Bessières, P., Bolotin, A., Borchert, S.,
Borries, R., Boursier, L., Brans, A., Brana, P., Brigneau, S.C.,
Bron, S., Brouillet, S., Bruschi, C.V., Caldwell, B., Capuano, V.,
Carter, N.M., Choi, S.K., Codani, J.J., Comerford, I.F., Cummings, N.J.,
Daniel, R.A., Denizot, F., Devine, K.M., Dusterhoft, A., Ehrlich, S.D.,
Emmerson, P.T., Enlihan, K.D., Errington, J., Fabret, C., Ferrari, E.,
Fougeret, D., Fritz, C., Fujita, M., Fujita, Y., Funa, S., Gallizzi, A.,
Galleron, N., Ghim, S.Y., Glaser, P., Goffeau, A., Golightly, E.J.,
Grandi, G., Guisepi, G., Guy, B.J., Haga, K., Halech, J., Harwood, C.R.,
Henaut, A., Hilbert, H., Holsappel, S., Hosono, S., Hullo, M.F.,
Itaya, M., Jones, L., Joris, B., Karamata, D., Kasahara, Y.,
Klaert, Blanchard, M., Klein, C., Kobayashi, Y., Koetter, P.,
Koningstein, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A.,
Larindois, S., Lauber, J., Lazarevic, V., Lee, S.M., Levine, R.P.,
Masuda, S., Mausel, C., Medigue, C., Medina, N., Mellado, R.P.,
Mizuno, M., Mostl, D., Nakai, S., Noback, M., Noone, D., O'Reilly, M.,
Ogawa, K., Ogiwara, A., Oudega, B., Park, S.H., Parro, V., Pohl, T.M.,

NII LQGEAASGVVWSPVSNIGLQVEVDPLPLEQTEPEEFLENETPKARSHOVLNLT
 NITLQGEAASGVVWSPVSNIGLQVEVDPLPLEQTEPEEFLENETPKARSHOVLNLT
 LGTEFAOTETPEKHNMSGGRVAKIKMAVILIEEDQVILIDPTPHILIDRSEOLLENTLS
 OYSGILLAVSHDRYKLETKTNSKIVYSNNGTEQLDNDVSENRREELRIKLETEROE
 VLGRKSEMTPNDRKTEKELDOATNELTKRIKELDHQKCD
 complement (5428. .5450)
 /gene="ydgf"
 complement (5428. .7089)
 /gene="ydgf"
 complement (5678. .5703)
 /gene="ydgf"
 complement (5713. .7089)
 /gene="ydgf"
 /function="unknown"
 /note="similar to amino acid ABC transporter (permease)"
 /codon_start=1
 /transl_table=11
 /protein_id="CAB12381.1"
 /db_xref="GI:2632875"
 /db_xref="SWISS-PROT:P96704"
 /translation="MTDDTKTNNIQDTLQRCIKNRHQLIAIGAIIGTGLFLGSGKS
 IHWGSPSLTVMYMGIGICFLILMRSGELLNSLNVHSPVDEVDYALGMAATITQMT
 IYFPCWISLAMDIDLVAGTQVLEPGVPMVGLALILILINLATVTKLEGELEME
 ALIKVIALITLIVLIGELNMFKEFSNSGVSTPMLMSGHFLPNCMGHFLTSOMVY
 APLVGLIGLTVIGELNPEKVIPIKANNIPRYLLFLYIGALVYKSIYMDITINPDS
 DEVDYFAVAVGIVGASITINFEVLTSPASACNSAVETSRVYSLADNHAPESMARKLT

Query Match	48.6%;	Score 807.4;	DB 2;	Length 213190;
Best Local Similarity	68.0%;	Pred. No. 2.6e-143;		
Matches 1126;	Conservative	0;	Mismatches 531;	Indels 0;
			Gaps	0;

OY	1	GAATTCGGCTTATATTGGCAAAAGAAATCAATTTTCACACAGATGGCGGTGCTGCATATG	60
Db	49236	GAGGTCCTTGTAAACATGCGCAAAAGAAATTAAGTTTAGTAAGAGATGCGCCGGCATACG	49295
OY	61	TCCCGCGAGTGAATATGTTAGACATACCGTCAAAAGTAACGGTTGGTCTTAAAGGCGCA	120
Db	49296	TTCGCGGTCATCATCACTTGGTATGCTGTAAAGTAATCTTAGAGCCAAAGACGCA	49355
OY	121	ATGTGTTCTTGA AAAAGCTTTTGGTCTCCCTTAATTCTAATACGGGGTAACCATTG	180
Db	49356	ACGTGGTCTTAGAGAAAATTCGGTTCTCCGTTATTCACAAATGCGGTGTACATCG	49415
OY	181	CTAAAGAGATCGAATTAGAAAGATCATTTTGGAAACATGGGAGCAAAATTGCTGTGAAG	240
Db	49416	CTAAAGAAATCGAGCTAGAAAGACGGGTTTGAAACATGGGTCTAAAGTTGTTGGTGAAG	49475
OY	241	TGGCTCTTAAACCAATGATATTGCTGGTGGATGGAGCAGCTACTGCACAGTTTGGACAC	300
Db	49476	TAGCCAGCAAAACCAACGCGTTGGCCGGTGAGGATCAACAACATCGATTCCTTGGCCG	49535
OY	301	AAGCATTTGTTATGAAGAGCTAAAAAATGTACACGAGGTGCTATATCCAAATTGGATATCC	360
Db	49536	AAGCATATGATCGGTGAAGGCCCTTAAAAACGTAAACGACGCGCTTAACCTGTAGGCGTGC	49595
OY	361	GTCGAGGCATTGAAACAGCACAGACAGACAGCTGTTGAAGCCTTGAAGCCATTGCTCAAC	420
Db	49596	GTTAAAGGATGAGAACAGCTGTAGCGGTTGTGCATCGAAAACTTAAAGAAATTTCTAAGC	49655
OY	421	CTGATCTGCGCAAGGAAAGCTATTGCTCTCAGAGTGGCTGCAATATCATCAGCTCTGAAAAAG	480
Db	49656	CAATCGAAGGCAAAAGTATATCCCTCTCAGGTTGCTGCACATCTCTCTGATTTGAGGAAG	49715
OY	481	TTGAGAGATATATCTCAGAGAGCTATGAGACGCTGTGGCAACAGATGTTGATATTACATCG	540
Db	49716	TGCGAAGCTTATTCCTCTGAAGCAATGGAGCCGCTTAGGAAACGACGCGTTATACAAATCG	49775
OY	541	AAGAAATCTCGAGATATGGAACAGACAACTTGAAGTGTGAAGCATGCATATTGACCGTG	600
Db	49776	AAGAGTCTAAAGGCTTCCACAACTGAGCTTGAAGTTGTTGAAGGTATGCAATTCGACCGG	49835
OY	601	GTTACCTGTCACATACATGTCACACAGACATGAAAAAATGGTTGCAACGCTTGAAAAAC	660
Db	49836	CATATGCGCTTCTTACATGTTGAATCTGACTGTGATATAGATGAACCGGCTTCTTGCAATC	49895

QY	661	CATTTCCTTAATACAGGATATAAAAAAGTGTCAAACTTCGACATATTTTGGCCACTCTCTTG	720
Db	49896	CTTACATCTTAATACAGACGACAAAAAATACAAACACTTTACGAAATCTTCTCTGTCTTG	49955
QY	721	ACGAAGTTCWTAAAAACAACGCTCCATTACTGATATTTGACAGATGATCGATGTGAAG	780
Db	49956	ACCAGGTTCWTACGACGAAGCAACACTTCTCTCTGATCTGATGAGATGTGAAGGCGAAG	50015
QY	781	CACCTCCACCCCTTGCTTGAACCAAGATTCGTGTACTTTCATATGTGTGTCTGTCAAG	840
Db	50016	CACCTTGCTACACTTGTGTGGAACAAACCTTCCGGGCACTTCAACGACGAGGCTGTAAAG	50075
QY	841	CGCCAGGATTTTGGTATCTGTCGTAAGATATCTTGAACACATTTGCTATCTTGACAGTG	900
Db	50076	CTCCGTGTTTGGGTATCGCCCTTAAGCAATGCTTGAGACATCTGCTGTCTTACTTGCGG	50135
QY	901	GTACAGTGATTAACAGAGATCAGACCTTGAAGCTTAAAGATGCTACAAATGACAGCCCTTG	960
Db	50136	GAGAAATCATACAGAAAGATCTTGGCCCTTGACCTGAATCTACCTCAATCGCTCAATTGG	50195
QY	961	GACAGGCTGCTAAGATTACAGTTGATTAAGATAGCAGCTAATTGTTGAAGTTACAGAA	1020
Db	50196	GACGGCGCTTCTAAATGTGCTTACTATAAGAAACACAACAATCGTTGAAGCGCTGGCG	50255
QY	1021	GTTCGAAGCTATTTGCTTAACCGTATTTGCACTGATTAATAATGCAATTCGAAACAACAAT	1080
Db	50256	AAACGACACAAATTTCTCGCCCTGTGACTCAAAATCCGCGCTTAAGTGGAGAAACAACTT	50315
QY	1081	CTGACTTTGACCGTGAAAAAATCTACAAGAACGTTTGGCGAATTTAGCTGGTGGTGAAGTG	1140
Db	50316	CTGAAATTCGACAGAAAAAATTAACAAGACGCTTGCTGCTAACTCTCGGCGCGTGAAGTG	50375
QY	1141	TTATGAAGTNGAGACTCCACAGACAGACAGCTTTAAAGAAATGAACCTTGCATTGAAG	1200
Db	50376	TCATCAAGGTGCGGCTGAGTGAACCTGACCTGAAGAGAGCGTAAACCTTCCACTCGAAG	50435
QY	1201	ATGCTCTAAATGCTCTACAGGTGACGCGTTGAAGAGATCTCTGCTGAGTGGTGAACAG	1260
Db	50436	ACGCTTGAACTCAACCTGGCGAGCTGTTGAAGAAAGGATCTATCCGCTGGTGGTACAG	50495
QY	1261	CACCTTATTACGCTTATTGAAAAAGTAGACGCTCTTGACCTTGAGGGCGATGCTACTG	1320
Db	50496	CGCTTGTAAAGCTATATATAACAAGTCGCTGCACTGAAGCTGAAGGCGATGCTCAAAACAG	50555
QY	1321	GACCTAACATATGCTGTGTGTCTGAAGAGCGCTGATGCTCAATTCCTTTAATGCTG	1380
Db	50556	GTAATCAATCTGCTGTGCGGCTTGAAGAGCCATCGCTCAAAATTCGCACACAACGCTG	50615
QY	1381	GGTACGACGCTCCCTAGTATTGACAAGTGAAGAAACAGCCCTCAGAGAACAGGATTTA	1440
Db	50616	GCCTTGAAGGATCTGTATCTGTTGAGGCCCTCAAAAAGGAAGAAATTCGCGTAGCTTCA	50675
QY	1441	ATGCTGCAACAGGTGAGTGGGTTGATATGATTAATAACAGAAATCAATGACCCGTGCAAG	1500
Db	50676	ACGCTGCAACAGGTGCAAAATGATTAACAATGATGCAAAAAAGTATCGTTGACCCAAACAAAG	50735
QY	1501	TAAACGATTCAGCGCTTAAAAATGCAAGCTTCTGTAGCTAGTCTTATTTTGACACAGAA	1560
Db	50736	TTACACGCTGACGCTTCAAAAACGCTGCTGTGATGCTGAATGTTCTTAACACACAGAA	50795
QY	1561	CAGTGTGCTTAATTAACCTGAACAGCTAGCCGACGCGCCAGACATGCGACAGSTAAAG	1620
Db	50796	CTGTGTGTGCTGACACAGCGACGAAGAAACGSGTGGCGGCGCAGGAATGCTCTGATATGGCG	50855
QY	1621	ATCCAGGAATGATGGGTGGATGGGCGGATTAAGCCGA	1657
Db	50856	GCATGCGCGGCTGATGGGTGAATGTATTAATAAGGA	50892
RESULT	10		
LOCUS	BACROEL	2190 bp	DNA
	BACROEL		BCF
			03-FEB-1999

RESULT	10			
BACROEL				
LOCUS	BACROEL	2190 bp	DNA	BCT
				03-FEB-1999

Db	1515	CTGAATTGCACAGAGAAAATTTACAGAGGCTTGTCTAACTTGTGTGGCGGCGTACTG	1574
OY	1141	TTTATCAAAATGAGAGCTCCACACAGACAGCTTTAAAGAAATGAAACTTTCGATTTGAGG	1200
Db	1575	TCATCAAAAGTCGGCGCTGCGCACTGAAACCTGACTGAAAGAGCGTAAACTTCGCATCGAAG	1634
OY	1201	ATGCTCTAAATGCTACACGTCGACGCCGTTGAGAAAGTATCGTTGCTGTGGTGGACAG	1260
Db	1635	ACGGCTTGAACCTCAACTCGGCGACAGCTTGGAACAAAGCATCGTATCCGATCGGTGGTACAG	1694
OY	1261	CACTTTATACGGTTATTTGAAAAAGTATGACAGCTTTAGCTTGAAGGCGATGATGCTACTG	1320
Db	1695	CGCTTTGAAACGTTATATACAAAGTGCCTGCTACGTTTAACTAAAGGCGATGCTCAAAAG	1754
OY	1321	GACGTAACTTTTGCTTCGTGCTCTAGAAGAGCCTGTACGTCAAAATTCCTTTAAATGCTG	1380
Db	1755	GTATCAACATCTGCTCTTCGGCGGCTTGAAGAGCCATCGGTCAAAATCCACACAAACGCTG	1814
OY	1381	GGTACGAAAGCTCCGTTAGTTATTGACAAAGTTGAAAAACAGCCCTCAGAGAACAGATTTA	1440
Db	1815	GCCTTGAAGCATCTGTCATCTGTTGACCGCCTCAAAAACGAAGAAATCGCGTAGCTTCA	1874
OY	1441	ATGCTGCAACAGAGTGGGTTGATATGATTTAAACAGAGATTCATTTAGCCTGTCAAG	1500
Db	1875	ACGCTGACACTGGGGAATGGGTAAACATATGCAAAAGAAAGTATCGTTGACCACCAACAAAG	1934
OY	1501	TAAACAGCATCAGCGCTTCAAAATGACAGCTTCGTAGCTAGTCTTATTTTGGACAACAGAG	1560
Db	1935	TTTACAGCCTAGCTCTTCAAAACGCTGCGCTGTAGCTCAATGTTCTTTAAACACAGAG	1994
OY	1561	CAGTTGTTGCTAATAAACCCTGAACACAGCTACGCGACGCGACCAATGCGACAGGTATGG	1620
Db	1995	CTGTTTGCGCTGAACAAGCCAGAAAGGATGGCGCGCCAGAGATGCGTGAATTGCGCG	2054
OY	1621	ATCCAGAAATGATGGTGGGATGGGCGGATTAAGCCGATTT	1660
Db	2055	GCATGGGCGGATATGGGTGAGATATGATTAATAAGATATT	2094

Db 1084 GGCCTGCTTAAGGTGTTGTGACAAAAGACACACACATGCTGACGAAACGGAG 1143
 QY 1021 GTTCGAAGCATATGCTAACCGTATTCGATTAATTCGAATTAAGAACCAACTT 1080
 Db 1144 ATGAGCGGAATCTCTACCGTGTATTACCAATCAAGCCCAATCGAAGAACACTT 1203
 QY 1081 CTGACTTGGACCGTGAAGAACTACAGAACGTTTGGCGAAATTAAGCTGGTGTAGCTG 1140
 Db 1204 CTGATTCGATTAAGAAACCTCAAGAGCCCTAGCTAACTAGCTGGCGTGGCTG 1263
 QY 1141 TTATCAAGTGAAGCTCCACAGACAGACCTTTAAAGAAATGAATTCGCAATGAGG 1200
 Db 1264 TCGTAAAGTGGTGGCGCTACTAACTGAATGAAGAGCCGTAACCTTCGCAATGAGG 1323
 QY 1201 ATGCTTAATGCTACAGCTGACCGCTTGAAGAGATGCTGCTGGTGGAGAC 1260
 Db 1324 AGCGCTTAATCTCAACAGCTGACCGGTTGAAGAGGAAATGTTGCTGGTGGTACAG 1383
 QY 1261 CACTTATTCAGGTTATGAAAAGTAGCAGCTCTTGACCTTGAGCGCGATGATGCTACTG 1320
 Db 1384 CCTGCTTAAGCTCTGCTGACCTGCGCAGCATTAAGTTGAAGGTGACGACGACAG 1443
 QY 1321 GACGTAACTGCTGCTGCTCTAGAGAGCTGTACGTCGAATTCCTTTAAATGCTG 1380
 Db 1444 GTGTAAACATCTGCTGCTGCTCGCAGAGCAGCTGCTGAATCTCTGAAAACGCTG 1503
 QY 1381 GGTACGAAGCTCGGTATGTTATGACAAGTGAAGAAACAGCCGCGACAGACGATTTA 1440
 Db 1504 GACTTGAAGGCTCTATATCATCTCTGACCCCTGAAGGTGAAGAGTGGCATCGCTTTA 1563
 QY 1441 ATGCTGAACAGGTAGTGGTGTATGATTAATAACAGAAATCATGACCTGCTGAAG 1500
 Db 1564 ACGGAGTACTGGGGAATAGTGTATATG-TTGAAGCAGGAATGTTGACCAACGAAAG 1622
 QY 1501 TAACACGATCAGCCCTTCAAAATGCAAGCTTGTAGCTATGTTATTTTGAACAGAG 1560
 Db 1623 TAACACGCTCCGCCCTTCAAAATGCAAGCTGTATGCTGCTATGCTTAACACAGAG 1682
 QY 1561 CAGTTGTTGTAATAAACCTGACAG 1587
 Db 1683 CTGTTGTGCTGACCTTCCAGAGAG 1709

RESULT 12
 AP001508 296950 bp DNA BCT 03-AUG-2000
 LOCUS Bacillus halodurans genomic DNA, section 2/14.
 DEFINITION AP001508 BA000004
 ACCSSION AP001508.1 GI:10172890
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Bacillus halodurans DNA.
 Bacteria, Firmicutes; Bacillus/Clostridium group;
 Bacillus/Staphylococcus group; Bacillus.
 REFERENCE
 1 (sites)
 AUTHORS Takami,H., Nakasone,K., Hirama,C., Takaki,Y., Masui,N., Fujii,F.,
 Nakamura,Y. and Inoue,A.
 TITLE An improved physical and genetic map of the genome of alkaliphilic
 Bacillus sp. C-125
 JOURNAL Extremophiles 3 (1), 21-28 (1999)
 MEDLINE
 99184645
 REFERENCE
 2 (sites)
 AUTHORS Takami,H., Nakasone,K., Ogasawara,N., Hirama,C., Nakamura,Y.,
 Masui,N., Fujii,F., Takaki,Y., Inoue,A. and Horikoshi,K.
 TITLE Sequencing of three lambda clones from the genome of alkaliphilic
 Bacillus sp. strain C-125
 JOURNAL Extremophiles 3 (1), 29-34 (1999)
 MEDLINE
 99184646
 REFERENCE
 3 (sites)
 AUTHORS Takami,H., Takaki,Y., Nakasone,K., Hirama,C., Inoue,A. and
 Horikoshi,K.
 TITLE Sequence analysis of a 32-kb region including the major ribosomal

JOURNAL protein gene clusters from alkaliphilic Bacillus sp. strain C-125
 MEDLINE Biosci. Biotechnol. Biochem. 63 (2), 452-455 (1999)
 99209008
 REFERENCE
 4 (sites)
 AUTHORS Takami,H., Masui,N., Nakasone,K. and Horikoshi,K.
 TITLE Replication origin region of the chromosome of alkaliphilic
 Bacillus halodurans C-125
 JOURNAL Biosci. Biotechnol. Biochem. 63 (6), 1134-1137 (1999)
 MEDLINE 99356711
 REFERENCE
 5 (sites)
 AUTHORS Takami,H., Takaki,Y., Nakasone,K., Sakiyama,T., Maeno,G.,
 Sasaki,R., Hirama,C., Fujii,F. and Masui,N.
 TITLE Genetic analysis of the chromosome of alkaliphilic Bacillus
 halodurans C-125
 JOURNAL Extremophiles 3 (3), 227-233 (1999)
 MEDLINE 99411980
 REFERENCE
 6 (sites)
 AUTHORS Takami,H.
 TITLE Genome analysis of facultatively alkaliphilic Bacillus halodurans
 C-125
 JOURNAL (in) Extremophiles in deep-sea environments (Ed.);
 HORIKOSHI,K. TSUTSI:
 : 249-284; Springer-Verlag (1999)
 REFERENCE
 7 (sites)
 AUTHORS Takami,H. and Horikoshi,K.
 TITLE Identification of facultatively alkaliphilic Bacillus sp. C-125
 to Bacillus halodurans
 JOURNAL Biosci. Biotechnol. Biochem. 63, 943-945 (1999)
 REFERENCE
 8 (sites)
 AUTHORS Takami,H. and Horikoshi,K.
 TITLE Analysis of the genome of an alkaliphilic Bacillus strain from an
 industrial point of view
 JOURNAL Extremophiles 4 (2), 99-108 (2000)
 MEDLINE 20263314
 REFERENCE
 9 (bases 1 to 296950)
 AUTHORS Takami,H., Nakasone,K. and Takaki,Y.
 TITLE Direct Submission
 JOURNAL Submitted (22-MAR-2000) to the DDB/EMBL/Genbank databases. Hideo
 Takami, Japan Marine Science and Technology Center, Deep-sea
 Microorganisms Research Group; 2-15 Natsushima, Yokosuka, Kanagawa
 237-0061, Japan (E-mail: takami@jamstec.go.jp,
 URL: http://www.jamstec.go.jp/jamstec-e/dio/DEEPSAR/Research.html,
 Tel:81-468-67-3895, Fax:81-468-66-6364)
 FEATURES
 source
 1. 296950
 Location/Qualifiers
 /organism="Bacillus halodurans"
 /db_xref="taxon:8665"
 /note="alkaliphile"
 complement(427..615)
 /gene="BH0278"
 complement(427..615)
 /gene="BH0278"
 /note="unknown"
 /codon_start=1
 /transl_table=1
 /protein_id="BAB03997.1"
 /db_xref="GI:10172891"
 /translation="MRKMGTLVLASIFVLVACNADENETESVDQNAEVEYTEDELE
 QDDEFEDEVVEETDDEE"
 798..1163
 /gene="BH0279"
 798..1163
 /gene="BH0279"
 /note="unknown"
 /codon_start=1
 /transl_table=1
 /protein_id="BAB03998.1"
 /db_xref="GI:10172892"
 /translation="MKKIFSVAAAFALASOALAAPETVSEQGVSNADYTPPE
 TVDAQSVWIMSGRITNTPTKTAIYNTYHRDSPHNSAALQNDRAVSLFGGGAFTT
 YASIRGNKNLPATFIRFTY"
 1334..3475
 /gene="BH0280"


```
CDS      1334..3475
         /gene="BH0280"
         /note="unknown"
         /codon_start=1
         /transl_table=11
         /protein_id="BAB03999.1"
         /db_xref="GI:10172893"
         /translation="MAHSAADAFVAVITPRIDLLGPDIVLSVEEAATHQVNIIR
         DVOLGEGDKIALTTFALLESPLLEAIPLRGSRLLSEDTREDGLILSRPTKEDY
         BEIKGTIQLTSQTHRTVRPLHQPNVSPVATYQVEQAGTVKQEQSATLADFNHY
         LEDYVELEPLPADPAVQTEPSHVSSTFDPPEPKTLAWTIMGITGLIMFYIMRKG
         TVAIYMHGSEFLKIRRLIGNAGVPAVYLLFPLFALEGMQPRILEAFAMWGS
         MPILLATISPLEFVSRSRLSESTKNTKTEFLIVSQAKVAVALISVYVLLDIA
         DLTITVEVRNSMTDMAFGKDGFEYTPHSIGDQEEHNOESGLMOTDISLTSFLFVL
         NAGALYDAVLYEEEMRNATEYSQEHIRSRIVNPNYLQAPLVDENNKRIALQEE
         TNRVLVPEQYQDKQEDILAYEKEDYAYATMPEYQAESYSHLOEQTFPIIMLASSQ
         VEFNLVYFPPENGNLIVDPILQVLETNNSYPMRDEPFGSVGNPLKPKLLDONTAKT
         YEHYLPRLDLDDNAKHLVYNEQAKDISLQALILDLTILFVLTATLAFMIVQ
         SSHLEFQHKRRFLERLFGHSLFRAYRNVLPTLTATWYIIGIALIKRHSQYLYAI
         VLVLFTEVAVTSVLMRLQKNKVSVNGE"
         /gene="BH0281"
         /gene="BH0281"
         /codon_start=1
         /transl_table=11
         /product="ABC transporter (ATP-binding protein)"
         /protein_id="BAB04000.1"
         /db_xref="GI:10172894"
         /translation="MAAAIVELTIDISKFSHSENVLEDFNLITDRCGMIAITGPGSSG
         TTILNIGLEKEDKGTVALFEKGNKRAVYSKQALQORRKAIVLPENFALMDEYISK
         NIDVPELINGSRTVAQOKKKEYLEKVFLENKLNKSVHSLSGEQQQVATARLLKPC
         DILLADEPTGSLDMANRNVILLDLCQLNKGMTIVITVTHDPEVADCDHYTL"
         /complement(4354..4641)
         /gene="BH0282"
         /complement(4354..4641)
         /note="unknown conserved protein in B. subtilis"
         /codon_start=1
         /transl_table=11
         /protein_id="BAB04001.1"
         /db_xref="GI:10172895"
         /translation="MKIEHVAIWVDLEAMKTFYTYTFNGHANSKYHNKDKPESTFL
         ITDGGARLEIKQSGIDKPDQDDRMGMHIAVSLGSEAVNQTGRLHNDG"
         /complement(4740..5537)
         /gene="BH0283"
         /complement(4740..5537)
         /note="unknown conserved protein"
         /codon_start=1
         /transl_table=11
         /protein_id="BAB04002.1"
         /db_xref="GI:10172896"
         /translation="MMTPPIYVDAFTNQAFKNPAACVLPSPRDDIMMOHVASPM
         LSETAPLHPYQGYSLRMFTPNTEVDLCHATLASHHILMELDHISAEPITPYTRSG
         ILTASRGEWILDFPSEQKQENVYPNLIDGLGLOPLVGNRNDYILEIDSEGL
         KLANPNEFLQIDIRGIIIVTSKSTSEYDFISRCFFPVGAVNEPVTSSACCLGPY
         MOEKLNKNEFLYQASKRGMILKIKLQHDREVLFGQAVIVLSEILL"
         /complement(7096
         /gene="BH0284"
         /complement(7096
         /gene="BH0284"
         /codon_start=1
         /transl_table=11
         /product="nicotinate-nucleotide--dimethylbenzimidazole
         phosphoribosyltransferase"
         /protein_id="BAB04003.1"
         /db_xref="GI:10172897"
         /translation="MOKVEMIGSINELDTKAQOOMEQHNLTLPGLSIGLIESIAI
         LKAGTGTQPEIDPATVYVADGVAAEYQASAEVYTOOMVHNFITGCAIANVLS
         RVSNAVOIVDIGEGLDLPGLINKNAGVGNMAAGNMAQSEERETIAIEGIVAO
         RAIONGARILLAGEMGIGNTTASSAMALAEVPPVQIYGFSTGSHSEKEKAVYIK
         RAIEVNPRAADPIDVLAKVGLEITAGIVLVGAAMRIPVLVDSFTTAVALTAVR
```

```
gene      IAPLCVHYLIASHQSYEPGHVCYNKLLGLTPLVNLRLRGEGSTIALPIYRSAIRI
         AHEMATFEQAGVSGAIEDVGRK"
         /complement(7227..7766)
         /gene="BH0285"
         /complement(7227..7766)
         /gene="BH0286"
         /note="unknown conserved protein in others"
         /codon_start=1
         /transl_table=11
         /protein_id="BAB04004.1"
         /db_xref="GI:10172898"
         /translation="MKSIYVIRCEAGQPPRESPLNKLKQALDAAFDNKIVRI
         ISSPYRAIQTOTPLAKKANVELINSQTERLYSLNLSDFEKLRTIFEDDELFE
         GESSQDAVKRIILNVEDSFTEFENTLIVTGHNLMLLNHNKFEFGDEWLNSNP
         DYLAKTDNNNETSERLWS"
         /gene="BH0286"
         /gene="BH0286"
         /codon_start=1
         /transl_table=11
         /product="ABC transporter (ATP-binding protein)"
         /protein_id="BAB04005.1"
         /db_xref="GI:10172899"
         /translation="MASHGKLEVDLQKNYKNNLTALQVNFILPGEFLTGM
         PGSQKTYLLNCIATIKPTSGVILNHNISFSPKDLAKRGSRIGYLFQDFEILD
         NLTGQRNILLPLAIGHVDFTKARKVIDELIASILDIADILHKRPSQMSGQOKORVAAAR

Query Match      47.3%  Score 786.4;  DB 2;  length 296950;
Best Local Similarity 67.3%  Pred. No. 2.2e-139;
Matches 111;  Conservative 0;  Mismatches 541;  Indels 0;  Gaps 0;

QY      4  TTGGGCTTCATATGGCAAAAAGAAATCAAAATTTTCAGCAGATGGCGTGCATGGATGGTG 63
         ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 295025  TTGAGATTAACATGGCAAAAAGAAATTAAGTTTACGGAAGACGACGCTGATGCTTC 295084

QY      64  GCGGAGTTGATATGTTAGCGATACCGTCAAAGTAACGCTTGCTTAAAGGCGCAATG 123
         ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 295085  GTGGTGATGATAAAGCTTGTCAGACGACATAAAGTAAAGCTTGAGACCAAAAAGCTGTAACG 295144

QY      124  TTGTCCTTGAAAAAGCTTTGTTCCCTTAACTATGATGACGGGGTATACCATGGCTTA 183
         ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 295145  TCGTTCTTGAAAAAGAAATTTGGTTCTTCQACTATTACAAATGACGGGTAAACATGCAATGCAA 295204

QY      184  AAGAGATCGAATTAGAACATCATTTGAAACATGGAGGACCAAAATGCTGTGAAGTGG 243
         ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 295205  AGGAAATCGAATTAGAACATGCGCTTGTAAGAACATGGGTGCAAAAGCTTGTGCAAGAGTGG 295264

QY      244  CTTCTAANAACCATGATATTGCTGTGTGATGGAGGACTACTGCAACAGTTTGCACACAG 303
         ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 295265  CAAGCAAGCAAAATGACATTTGCTGGGACGGTACAAACACGACGACGCTCTTGACACAG 295324

QY      304  CCATTTGTCATGAAGGCTTAAAAATGTGACACAGGTGCTAATTCGAATGTGATCCGTC 363
         ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 295325  CGATGATTCGCGAAGGCTTGAAAACGTCACGTCGTGGCGGAACCCCAATGTTATTCGTA 295384

QY      364  GAGGCAATTGAACAGCAACAGCAACGCTGTGAAGCTTGAAGCCATTCCTCAACCTG 423
         ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 295385  AAGGAATCGAAAAAGGACGACGATGACGCTTGAAGACTTCTTAATGCTTAAGCCAA 295444

QY      424  TATCTGGCAAGAGCAAGCTATTGCTCAGGTGCTGCAGTATCATCAGCTGTGAANAAGTTG 483
         ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 295445  TCGAAGGCAAAAGATTCAATGCGTCAAGTTGCGGCGGATTTCTTCTGTCGAGATGAAGTAG 295504

QY      484  GAGAGTATATCAGCAAGCATATGACGCTGTGGGCAACGATGCTGTGATATACATGGAAG 543
         ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 295505  GGAATAATCATTTGCTGAAGCATATGACGCGGTAGCAACGCGGCTTATTATGACATTGAAG 295564

QY      544  AATCTGAGCTATGCAAAACGAACTGGAAGTGTGTAAGGCAAGCAATTTGACCGTGTT 603
         ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 295565  AGCTTAAGGATTTCTTACAGACACTAGAAAGTGTGAAGTATGCAATTTTGACCGGGCT 295624

QY      604  ACCTGTCTCATCATGATGCTCAGACAAATGAAAAAATGTTGTGACGACCTGAAACCACAT 663
```


Db 295625 ATGGCTTCCTTACATGCTACAGACTCTACAGATGAGAGCGCTTCTTGACAAACCTT 295684
 QY 664 TTATCTTAATACAGAGATAAAAGTGTCAAAACATTCAGACATTTGGCCACTACTACTAGG 723
 Db 295685 ATGTCTTGATCAGAGATNAGAGATTTCTTACATCCAGAGAGTTTACAGCTTTTACG 295744
 QY 724 AAGTCTTAAACCAACCGCTCATTTACTTATTTGAGATGATGTGATGTGAGAACAC 783
 Db 295745 AAGTTGTACAACAAGGACGACCAATCTTATTCATTTGCTGAGAGATGTGAGAGTTGAC 295804
 QY 784 TTCCAAACCTTTGTCTTAAACAGATTTGTGTACTTCAATGTGTGTCTGTCAAAACGC 843
 Db 295805 TTGCAACACTTGTGTGTAACAACTTGTGTGAACTTTAACGAGTGTGCGTTAAACCTC 295864
 QY 844 CAGGATTTGGATCGTGTAAAGCTATGCTTGAAGACATTTGCTATCTTGTACAGGTGTA 903
 Db 295865 CTGGATTCGGTATCGTGTGTAAGCAATGCTTGAAGCATTTGCGATCTTACTGCGGTG 295924
 QY 904 CAGTGAATTACAGAGATCTAGGACTTGAATTAAGATGCTTACAAATGACAGCCCTTGAC 963
 Db 295925 AAGTATCACAGAAAGCTTAGGGCTGACTTGAATCTGCAACATCACAACAGCTTGCTC 295984
 QY 964 AGGCTGTAGATTTACGTTGATTAAGATAGCAGCTAATTTGTTGAAGTTTCAGAAATT 1023
 Db 295985 GCCCAAGTAAAGTCTGTGTACGAAAGAACACAGCATTTGTGAAGTGTGCGGCAAA 296044
 QY 1024 CAGAGCTATTTGCTAACCGTATTCATGATTAATGCAATTAAGCAATTAAGCAATCTTG 1083
 Db 296045 GCGTAAATTCGACGCTCGGCTTATCAATCAAGACACAAATCGAAGAAACAACATCTT 296104
 QY 1084 ACTTTGACCGTGAAGAACTACAGAACAGTGTGGCAATTAAGTGTGTGTACTGTGA 1143
 Db 296105 ACTTCGATTAAGAAACTTCAAGAGCGCTTAGGCAAGCTTGTGCGGTGCTGCTGTC 296164
 QY 1144 TCAAGTAGAGCTCCACAGACAGACCTTTAAAGAAATGAACCTTGCATTTAGAGATG 1203
 Db 296165 TTAAAGCGGTGACGCTACCTGAATGAAGAGCCCAACTTCGATTTGAAGACG 296224
 QY 1204 CTCTAAATGCTACAGCTGACGCCGTTGAAGAGTATGTTGCTGCTGCTGAGACAGAC 1263
 Db 296225 CGTTGAACCTTACTCGCGCGCTGTAGAAAGAAAGATCGTTGCTGCTGCTGCTGCC 296284
 QY 1264 TTATACGTTATTAAGAAAGTACAGCTCTTGAAGTGAAGGCGATGATGCTACTGAC 1323
 Db 296285 TTGTGAACGTGATTAAGACAGTCTTGAAGTGTGCAAGAGTGCAGAGCAACAGGTG 296344
 QY 1324 GTACATTTGCTTCTGCTCTAGAGAGCTGTAGCTCAATTTGCTTTAATGCTGAGT 1383
 Db 296345 TGAACATCGTCTTCTGCTGCTGAAAGAGCCAGTTCTCAANTGCTCACAAACGAGT 296404
 QY 1384 ACGAAGCTCCGCTAGTATTAAGTGAAGTTGAAGAAACACCCCTGCGAGAAACAGATTATG 1443
 Db 296405 TTGAAGCTCTTGATGCTGTGAGCGCTCAAGAAAGAAAGAGCTTGCTGGTTTCAAG 296464
 QY 1444 CTGCACAGGTGAGTGGTGTATATGATTAACAGAGATCATTTGACCTGTCAAGTAA 1503
 Db 296465 CAGCAATCTGGAATGGTGAACATGCTTGAAGTGTGATCGTTGACCAACGAAGTAA 296524
 QY 1504 CAGATAGCGCTTCAAAATGACAGCTTCTGAGTACTTATTTTACAAACAGAGACAG 1563
 Db 296525 CTGCTTACGCTTACAGACAGCATCTGATCTGCTATGCTTCTCAACAAACAGAGCG 296584
 QY 1564 TTGTTGTATTAACCGTGAACAGCTACGCGACCGCCAGCAATGCCAGAGTATGATC 1623
 Db 296585 TGATCGTGTATTAAGCTTGAGAGAGAACAGAGGCGCGGAGGATCTGACATGTGGCGGAA 296644
 QY 1624 CAGGAATAGTGGTGGATGGCGGATAGCC 1655
 Db 296645 TGGGTGCGATGGCGGCGATGATTAATTAGGC 296676

RESULT 13

AF010281
 LOCUS AF010281 3128 bp DNA BCT 09-AUG-1997
 DEFINITION Lactobacillus zeae GroES (groES) and GroEL (groEL) genes, complete cds.
 ACCESSION AF010281
 VERSION AF010281.1 GI:2286186
 KEYWORDS Lactobacillus zeae.
 SOURCE Lactobacillus zeae.
 ORGANISM Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae; Lactobacillus.
 REFERENCE 1 (bases 1 to 3128)
 AUTHORS Murphy, C.M. and Chassy, B.M.
 TITLE Molecular Characterization of the Heat-Shock Regulated groESL operon of Lactobacillus zeae
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 3128)
 AUTHORS Murphy, C.M. and Chassy, B.M.
 TITLE Direct Submission
 JOURNAL Submitted (25-JUN-1997) Food Science and Human Nutrition, University of Illinois at Urbana-Champaign, 1302 W. Pennsylvania Avenue, Urbana, IL 61801, USA
 FEATURES
 source
 -35_signal
 -10_signal
 repeat_region
 RBS
 gene
 CDS
 gene
 CDS
 terminator
 BASE COUNT 913 a 659 c 789 g 767 t
 ORIGIN
 1. 3128
 /organism="Lactobacillus zeae"
 /strain="1025 (Plasmid-free derivative of ATCC393)"
 /db_xref="taxon:57037"
 735..740
 /evidence=experimental
 738..763
 /evidence=experimental
 775..801
 /note="controlling inverted repeat of chapone expression: CIRCE"
 /pt_type=inverted
 819..823
 /evidence=not_experimental
 831..1112
 /gene="groES"
 831..1112
 /gene="groES"
 /note="HSP10; molecular chapone"
 /codon_start=1
 /transl_table=11
 /product="groES"
 /protein_id="AAB6325.1"
 /db_xref="GI:2286187"
 /translation="MLKPLGDRVIVEVEEETGVGIVLANNKOKPOTGKVVAVGE
 GALTPEGRKLPMVVKVDVLYDKVAGSEVKEGODYLVLHERDMAIA"
 1149..2783
 /gene="groEL"
 1149..2783
 /gene="groEL"
 /note="HSP60; molecular chapone"
 /codon_start=1
 /transl_table=11
 /product="groEL"
 /protein_id="AAB6326.1"
 /db_xref="GI:2286188"
 /translation="MAKEIKFSEDAARMLRGVDOLANTVKTTLGKGRNVLDKSYG
 SPEITNDGVIAKSIDLDEHYENMGAKLVAEVAASKNDIAGDCTTATVLAOSIIEG
 MKNVTAGANVGTIRTEIEKATKADELHRIKSHKVGKKEIADVAVSSSTEVGSLI
 ADAMERVGHGVGTIESKGIIDRELSVBEQMOPDRGLSOUVYTDNDKMRADIDDPYI
 LITDKISNODILPILOETVOGKALLITADVACEALPTVLNKTIRGFNVVANKA
 PGFDRKROLDIATITIGSTVLSIDGLDKTKLEQIGRGRVTVTKNTTIVGA
 GSKDAIAERNIITKQIDPTTSDPREKLOERLAKLAGVAVKVGAAETELKERY
 RIDDALNATRAAIVEGIVVAGGTALVDVLPVAVALKEEGVOTGINIVLRALEPVRQ
 IADNAREGSIIVLEOKKEKOGVYNATDEMMAKSGIIDPTKTVRSALQMAASVA
 ALMLTEAVVADKPDNPANNANNAAGANPAAGMGMM"
 2807..2830
 /note="putative"

Query Match 46.5%; Score 773; DB 1; Length 3128;
 Best Local Similarity 67.2%; Pred. No. 1.8e-136;
 Matches 1094; Conservative 0; Mismatches 535; Indels 0; Gaps 0;

15 ATGGCAAAAGAAATCAAAATTTTCAGCAGATGCGGTGCTCCATGCTGGCGGAGTTGAT 74
 1149 ATGGCAAAAGAAATTAATTTCTGGAAGACGACGTGCTCAATGCTGGCGGAGTTGAT 1208

75 ATGTGAGATACCGTCAAGAGTAAGCTTGGTCTAAAGGGGCAATGGTGTCTTGA 134
 1209 CATTGGCAACACAGTTCAGACACCTTGAACCAAGCGGCGGACGTTGTTGGAC 1268

135 AAGCTTTTGTCTCCCTTAATTAATGACGGGGTAACCATTTGCTAAAGATCGAA 194
 1269 AAGAGCTATGGCTCACCGGAAATTAACCAATGACGGGTGTCAATTTGGAAAGTGCATTGAC 1328

195 TTAGAAGATCATTTTAAACATGGACCAAAATTTGGTGTCTGAAGTGGCTTCAAAACC 254
 1329 TTGGGAAGATCACACGAAACATGGGCGCCAGTGTGTCGCAAGTGTCTCAAGACA 1388

255 AATGATATTGCTGTGATGGGAGGACTACTGCAACAGTTTGAACAGGCAATTTGCAT 314
 1389 AATGACATTGCTGTGATGGTACCAACCGCAGCGTTTGGCAACAGTATTATTCGCC 1448

315 GAAGACTAAAAAATGTGACAGCAGGTGCTAATCAATTGATTCGTCGAGGCAATTTGA 374
 1449 GAAGGTATGAAAGCTTACAGCGGGTGGCTAATCTGTTGGCATTTGCAACAGGATTTGA 1508

375 ACAGCAACAGCAACAGCTGTGAGCCCTTGAAGCCATTGCTCAACCTGATCTGGCAAG 434
 1509 AAGGCACTAAGGCTCCCTTGGAGTAATGCACAGATTAAGCCAAAGTTAATGGTAAAG 1568

435 GAAGATATTGCTGAGTGGCTGCGATATCATACGCTCGAAAAGGATGAGAGTATATC 494
 1569 AAGAAATATGCGCAGGTGGCTGCGTTCTCTCAAAATACAGAGTGGTACTGATTT 1628

495 TCAGAGCTATGAGGCTGTGGGCAACGATGTGATTAACATCGAAGAAATTCGAGGT 554
 1629 GGTGACGCGCATGSAAGAAAGTTGGCCACGATGGTGTATTACATTAAGAAACCAAGG 1688

555 ATGGAACAGCAACTGTAAGTGTGAGAGCATGCAATTTGACCGGTGTTACCTGTCTCAA 614
 1689 ATTGACACTGAACCTCTCTGTTGTAAGGAGATGCACTTTGATGCGGCTATCTGACCCA 1748

615 TACATGGTCCAGACAGCAATGAAAAATGTTGACAGCTGGAAGAAACCATTTATCTAATC 674
 1749 TACATGGTACCGATATATATATAGATGAAAGCTGACCTTGACATCTTATATCTTGATC 1808

675 ACGGATAAAAAAGTGTCAAAATCCAGACATTTTGCACACTACTAGGAGAAAGTTCTTAA 734
 1809 ACCGACAAAAGATTTCCAAATTTTCAGGACATTTTGCCTGTGTACAGAAATCGTTCAA 1868

735 ACCAACCGTCCATTAATCTTATGACAGATGATGTGATGGTGAACACTTCCAAACCTT 794
 1869 CAAGGTAAAGCACTGTGATCACTTGTGACGAGCTGTGGTGAACACTTGCACACCTTA 1928

795 GTCTTAAACAGATGCTGCTACTTCAATGTTGTTGCTGCTCAAAACGCGAGAGTTGGT 854
 1929 GTTCTGAACAAAGTTGCTGTACCTTCAATGTTGTGCGTTAAAGCCCTTGCTTGGGT 1988

855 GATCGTGTAAAGCTATGCTTGAAGCATTTGCTATCTGACAGGTGTACAGTATGATTA 914
 1989 GATCGGCTGAAGGCTCACTTGAAGATTTGGTACTTGAAGTGTGTACCGGTGATCAAGT 2048

915 GAGGATCTAGGACTGATTAAGAAATGCTTACATGACAGCCCTTGGACAGGCTGTGAAG 974
 2049 TCTGACCTTGGCTGTGATTTAAAGACACCAAGCTTGAACAAATTTGGTGGCGGCGCAAG 2108

975 ATTACAGTGAATAAGATGACAGCAATTTGTTGAAGGTTTCAGGAGTTCAGAGCTATT 1034
 2109 GTTACGGTTACCAAGGACACACACAGATTGTTGATGGCGCTGCTTAAGCATGGCAATT 2168

QY 1035 GCTAACCGTATTCACATGATTAATGCAATWTAGAAAACAACAACTTCTGACTTTGACCGT 1094
 DB 2169 GCGAAGCGGTTAATATCATCAAGAAAGCAATTTGATGACGACGACGTGCTTGCACCGT 2228

QY 1095 GAAAACTACAGAACGCTTGGCGAAATTAAGTGTGTGTAGCTGTTATCAAGTAGGA 1154
 DB 2229 GAAAGGTGCAAGAGCGTTTGGCTAAGTTAGTGGCGGTGTTCGGTGTCAAGTTGGT 2288

QY 1155 GCTCCACAGAGACGCTTAAAGAAATGAACCTGCGATTGAGATGCTCAATATGCT 1214
 DB 2289 GCGGCTACTGAAGATTAATGAAGAAAGCAAGTATGCGATGCGAAGATGCTTGAAGCG 2348

QY 1215 ACACGTGACCGCTTGAAGAGGATATGCTGTGATGTGAGACAGCAGCTATTACGTT 1274
 DB 2349 ACCGTGCGCGGTTGAAAGAGTTAGCTTGTGGGCGGCTACAGCTCGTGTGATGATGT 2408

QY 1275 ATTGAAAAAGTACGACCTCTTGAGCTTTGAGGGCGATGATCTACTGAGCTAACATTTG 1334
 DB 2409 TTCCCTGCAAGTGTCTGCTTGAAGAAAGAGCGACGCTTCAACCGGATCAACATCTG 2468

QY 1335 CTTCGTGCTGTAGAGAGCGCTGTACGTCATTTAAATGCTGGTACGAAGGCTCC 1394
 DB 2469 CTACGTGCGCTTGAAGACCTGTTCCGCAATTTGCTGAGACGCGCGCAAGAGGTTCT 2528

QY 1395 GTAGTTATTGACAAGTTGAAAAACAGCCCTGACAGAAACAGATTTAATGCTGCAACAGT 1454
 DB 2529 GTCATGCTTGAACAGCTGAAGAAAGAAAGCAAGGCGGTTGTTCAATGCGGCTACTGAT 2588

QY 1455 GAGTGGCTTGTATGATTTAAACAGCAATCTTACCCCTGTCAAGTAACAGATACGCG 1514
 DB 2589 GAATGGGAAGATATGCTTAAGTACGATTAATTAATGACCAACCAAGTGAACCCCTTCTGCA 2648

QY 1515 CTTCAAATGACAGCTCTGTAGTACTTATTTTACAAACAAGCAGTGTGCTAAT 1574
 DB 2649 TTGCAAAAGCGCGCTTCCGTGCTGTGATGCTACGACGATGAAACCGTGTGCTGAT 2708

QY 1575 AAACCTGAACAGCTACGCGAGCGCCAGCAATGCCAGAGTATGATCCAGAGATGATG 1634
 DB 2709 AAGCCAGATCCAAATTAATATATATATGCTGCCGCTGTGCTTAAACCGAGAGGCTATG 2768

QY 1635 GGTGGGATG 1643
 DB 2769 GCGGCTATG 2777

RESULT 14
 AF214488
 LOCUS AF214488 2753 bp DNA BCT 13-JUN-2000
 DEFINITION Lactobacillus johnsonii groESL operon, complete sequence; and unknown gene.
 ACCESSION AF214488
 VERSION AF214488.1 GI:8489166
 KEYWORDS
 SOURCE
 ORGANISM
 Lactobacillus johnsonii.
 Lactobacillus johnsonii.
 Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae; Lactobacillus.

REFERENCE
 1 (bases 1 to 2753)
 Walker,D.C., Girgis,H.S. and Klaenhammer,T.R.
 The groESL chaperone operon of Lactobacillus johnsonii
 JOURNAL Appl. Environ. Microbiol. 65 (7), 3033-3041 (1999)
 MEDLINE 99318659
 PUBMED 10388700

REFERENCE
 2 (bases 1 to 2753)
 Walker,D.C., Girgis,H.S. and Klaenhammer,T.R.
 Direct Submission
 Submitted (09-DEC-1999) Food Science, North Carolina State
 University, Room 341 Schaub Hall, Raleigh, NC 27695-7624, USA

FEATURES
 source
 1..2753
 /organism="Lactobacillus johnsonii"
 /strain="VPI 11088"

```

CDS      /db_xref="taxon:33959"
          complement(179..331)
          /note="oriS"
          /codon_start=1
          /transl_table=11
          /product="unknown"
          /protein_id="AAF5591.1"
          /db_xref="GI:8489167"
          /translation="MTAFGAVLAFISFLDCIVIAFIKLTNDTNKILVEAVKRLIETLH
          DLKKK"
          652..2598
          /gene="groESL operon"
          652..936
          /gene="groES"
          652..936
          /note="chaperone protein"
          /codon_start=1
          /transl_table=11
          /product="GroES"
          /protein_id="AAF5592.1"
          /db_xref="GI:8489168"
          /translation="MLQPIGDRVIVVKKDEEEKVGVIVLASNAKEKRPQGEITAVN
          GKRNSDLIPMSVAKGETVFPDKYSTNLKTEGKTLVRESDLAVK"
          967..2598
          /gene="groEL"
          967..2598
          /note="chaperone protein"
          /codon_start=1
          /transl_table=11
          /product="GroEL"
          /protein_id="AAF5593.1"
          /db_xref="GI:8489169"
          /translation="MAKEIKFSENAHSILKGVLDLADYKTLTGPKRNVLEKSG
          APTITDGVITAKSIELENHFEWMAKLVSEAOKNDIDGDTTATVLTQAVREG
          MKNVTAGANPVGIRRGLETATKAVDELHKISHKSTKDIADVASVSASSTEYGNL
          RDAMEVYGHGVITIEESKIDTELSEVEMGDRGLSOYMTDNDKMDADNPYI
          LITDKKISNODILPLOTVOGKSLIILADVDGALPTVLNKRGENNVAVNA
          PGFDRKAMLEIDIALITGSTVSSDGLKDKPTIDOLGAKKVTYTKRSTVEGA
          GSKETAIERYDQIKKOJADTSPDRKIDIERLAKLAGAVAVIKVCAITTELKERY
          RIDALNATAAVEGIVAGGTALVDVMSIGGVKGDSEDETKVIMKLGAV
          RQIAENAGKGAVIDLHLEHDEPEVGNATNKNWNVKAGIIDPTKVTLSALONAS
          IAAILTEVNVVADAPEDDKQAPAAPNPQMGGM"
          901 a 458 c 571 g 823 t

BASE COUNT
ORIGIN
Query Match      46.4%; Score 771; DB 1; Length 2753;
Best Local Similarity 68.9%; Pred. No. 4.3e-136;
Matches 1088; Conservative 0; Mismatches 485; Indels 6; Gaps 2;

```

```

Db 1263 TCGTGAAGGTATGAAAGAACGTACTGCTGTGCAAAACCTGTGTGATTCGCGGGAT 1322
Oy 371 TGAACACGACACAGACAGCAGTGTGAAGCCTTGAACCCANTGCTCAACCTGTATCGG 430
Db 1323 TGAACCTCTACTAAGCAGCGTGTGACGAATTCACAGATTAAGCAACAGTAAGTAC 1382
Oy 431 CAAGAACCTATTTGCTCAGTGTGCTGCATATCATCAGCTCTGANAAGTTGGAGACTA 490
Db 1383 TAAGSACGAATTTGCTCAAGTGTCTCGTTCATCAGCTTCACTGAGAGTTGTAAGT 1442
Oy 491 TATCTCAGAACCTATGAGCGGTGGGCAACGATGTGTATACCATCGAANAATCTCG 550
Db 1443 AATCGGTACCCCAATGGAANAAGTTGTCAGCATGTGTATATCTATTCGAAAGTCAA 1502
Oy 551 AGTATGGAACACAGACCTTGAAGTGTGAAGGCATGCAATTTGACCGTGTACTCTGC 610
Db 1503 GGGTATGTATCTGAACTTCAAGTATGAGTGAAGATGGAATTCGATCGTGTACTATTC 1562
Oy 611 TCAATCATGTGTACAGACAAATGANAATGTTGACACCTTGAANAACCATTTATCTT 670
Db 1563 ACAATACATGTGTAAGTGAACAACAGATGGAAGCAGACTTAGCAACCATACATTTT 1622
Oy 671 AATCAGGATTAANAAGTGTCAAAATCCAAAGACATTTTGCACACTTCTGAGAGATTCT 730
Db 1623 GATTACTGACAAAGAAAGATTTCTAATATTCAGATATCTTGCATATTTACAANAATCGT 1682
Oy 731 TAAACCAACCCGTCATTTACTCATTTATTCAGATGATGTGATGTGAGACACTTCAC 790
Db 1683 TCAACAGGTAAGAGCTTATTAATCATCTGATGATGTGATGTGAGAGCTTCCAAC 1742
Oy 791 CCTGTCTTGAACAAGATTCGCGTACTTTCATATGTGTGTGTCGTCGAAGCGCCAGAT 850
Db 1743 TCTTGTTTTGAACAAGATTCCTGCTGACTTTTACCTTGTGTGTAAGGCTCCTGGCT 1802
Oy 851 TGTGTATGCTGTAAAGCTATGCTTGAAGACATGCTATCTTGAACGCTGTACAGTAT 910
Db 1803 TGTGTACCGCTGTAAAGCAATGCTGAGATATTCCTATCTTAACTGTGTGACTGAT 1862
Oy 911 TACAGAGATCTAGACCTTGAATTAANAAGTCATCAATGACAGCCCTTGACAGGCTGC 970
Db 1863 TTCTCAACACTTAGTCTTGAATTAAGGACACTAAGATTCATCAATTAAGTAAGGCTGG 1922
Oy 971 TAGATTCAGTGTAAAGTAGACAGCATTAAGTGTGAAGGTCAGGAAGTTCAGAAC 1030
Db 1923 CAAGTTCCTTAACCAAGGATTCACACTACTATGTGTGAAGTGTGTTCAAGGAAGC 1982
Oy 1031 TATTGCTAACCGTATTCACATGATTAATTCGCAATTAAGAACACACACTTCTGACTTTGA 1090
Db 1983 TATTGACAGAGGTGATGATCAATCAATCAAGACCAATTCGTACACTACTTCAAGCTTTGA 2042
Oy 1091 CCGTGAAAACTACAAGACGTTTGGCGAAATTAAGCTGTGTGTGATGCTGTATCAAGT 1150
Db 2043 CCGTGAAAACTACAAGACGTTTGGCTTACCTGTGTGTGTGTGTGTGTGTGTGTGTGT 2102
Oy 1151 AGGAGCTCAACAGAGACGTTTAAANAAGTAAGAACTTGGCAATGTGAGTGTCTTAA 1210
Db 2103 TGTGTCTGCTACTGAATGATTAAGGANAAGTAAGATCAAGATCGAAGCGCTTTGAA 2162
Oy 1211 TGTACACGTCAGCCGTTGAAGAGGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1270
Db 2163 CGCAACCGCTCCGCTGTGAAGAGGTTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2222
Oy 1271 GGTATTTGA---AAAAGTAGAGCTTGTGAGCTTGAGCGCATGATGC---TACTGACG 1324
Db 2223 TGTATGAAAGTCATCCAGGTACTGTAAAGGTGACACGAAGCAGAAAGCTGGGT 2282
Oy 1325 TAACATGTGCTGTGCTGTGAGAAAGCGTGTGCAATTTGCTTAAATGCTGGGFTA 1384
Db 2283 TAAAGTTGTTATGAAGCTTTAAGCGCTCTGTACGTGAATTCGTGAAGACGCTGGTAA 2342
Oy 1385 GGAAGGCTCCGTAGTTATGCAAGTTGANAAGACGCGCTCAGAGAACAGGATTTAATGC 1444

```

Db 2343 AGATGCTGCTTATCTTAGACCACTTAGACATGAAGACCAGAGTTGGTTACAATGC 2402

QY 1445 TGCACACAGTGAAGGCTGATGATGATTAACAAGACATCTTACCCTGTCACAACTAAC 1504

Db 2403 TGCACACTAACAGTGGGAAACATGCTTAAGCTGATATATGACCACTAACAGTAAAC 2462

QY 1505 ACATACAGGCTTCAAAATGACACTCTCTAGTACTGCTTATTTTGACACAAAGACAGT 1564

Db 2463 TCGTTCAGACCTTCAAAATGCTTCAATTCCTGCTTGTATTAATCACTAAGCTGT 2522

QY 1565 TGTTCCTAATAAAGCTGAA 1583

Db 2523 TGTTCCTGACGCTCCAGAA 2541

RESULT 15
AF269843/C 3885 bp DNA BCT 01-AUG-2000
LOCUS Staphylococcus epidermidis strain SRI clone step.1026c04 genomic
DEFINITION sequence.
ACCESSION AF269843
VERSION AF269843.1 GI:9623742
KEYWORDS
SOURCE Staphylococcus epidermidis.
ORGANISM Staphylococcus epidermidis.
REFERENCE 1 (bases 1 to 3885)
AUTHORS Kimmerly,W.J., Taylor,J.D., Nelsen,A.J., Godlevski,M.M.,
Rubino,M.A., Nelsen,F.J., Rivers,P.R., Tortorella-Miller,I.,
Listebee,S., Ashanti,C., Altschuler,G., Mammo,L., Shepherd,N.S.,
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
Furdon,P.J.
TITLE Transposon-mediated sequencing of the Staphylococcus epidermidis
genome
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3885)
AUTHORS Taylor,J.D., Kimmerly,W.J., Nelsen,A.J., Godlevski,M.M.,
Rubino,M.A., Nelsen,F.J., Rivers,P.R., Tortorella-Miller,I.,
Listebee,S., Ashanti,C., Altschuler,G., Mammo,L., Shepherd,N.S.,
Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
Furdon,P.J.
TITLE Direct Submission
JOURNAL Submitted (22-MAY-2000) Departments of Genomic Sciences and
Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
Drive, Research Triangle Park, North Carolina 27709-3398, USA
FEATURES
source 1..3885
/organism="Staphylococcus epidermidis"
/strain="SRI"
/db_xref="taxon:1282"
/clone="step.1026c04"
BASE COUNT 1061 a 739 c 560 g 1525 t
ORIGIN

Query Match 45.5%; Score 756.4; DB 1; Length 3885;
Best Local Similarity 67.5%; Pred. No. 2.3e-133;
Matches 1063; Conservative 0; Mismatches 511; Indels 0; Gaps 0;

Db 2691 GTTAGAAGATCATATGAGATATGGGTGCAAAATTAAGTCAGGAATTCGATTAAC 2632

QY 254 CAATGATATTCCTGCTGATGAGCAGACACTACTGCAACAGTTTGTACAAACCACTTGTCA 313

Db 2631 AATGGAATGCTGGGAGCGGTACMACTACAGACAGTTTGTAGCAATCAATGATTC 2572

QY 314 TGAAGACTTAAATAATGTGACAGAGGTGCTAATCCAAATGATTCGTCGAGCATGA 373

Db 2571 GGAAGGCTTAAAGAAATGTTCAAGGTGTCGAATCCGTAGGCTTAAGACAGATTTGA 2512

QY 374 AACAGACAGCAACAGCTGTGAAGCTTGAAGCATTCGTCACACTATCTGCGCA 433

Db 2511 CAAGAGAGTGCAGAGTGGCTATAGACGCTTCATGAAATTTCTCAAAAGGTGAATAA 2452

QY 434 GGAAGCTATTCCTCAGGTCGCTCAGATCATACGCTGAAAAAGTTGACAGATATAT 493

Db 2451 GAACGAGATGCGCCAGAGTTGAGCTATTTTCAGACAGATGAAGAAATGCTGCTACAT 2392

QY 494 CTCAGAGCTATGAGCGGTGTGGCAACGATGCTGATTAACATCGAAGATCTCGAG 553

Db 2391 TTTCTGAAGCAATGATTAAGTAGTACGATGCGCTTATCCTATTAAGAAATCAATGG 2332

QY 554 TATGGAACGAACTTGAAGTGGTTGAAGCATGCAATTTGACCGTGTACTGTCTCA 613

Db 2331 GTTTAATACGAATTAAGTAGTGAAGATGCAATTTGATCGGCTTATCAATCACC 2272

QY 614 ATACATGCTCACAGCAAAATGAATGTTGCGAGACCTTGAACCCATTTATCTTAAT 673

Db 2271 ATATATGATGATGACTGACTGATTAATGATAGCTGATTTGAACGTCATATATTAAGT 2212

QY 674 CACGATTAATAAAGTGTCAAAATCCAGACATTTTGCACATCTTGAGGAGTTCTTA 733

Db 2211 AACGATTAAGAAATTTTCATCATTCACAGATTTCTTCATTTAAGAACATGTTGCA 2152

QY 734 AACCAACGCTTCACTTACTATTTTCACATGATGTGATGGTGAAGCACTTCAACCC 793

Db 2151 GGTATGTCGACCAATTTTATTTGTTGCGGATGAAGTGAAGGAGTACATTAATAT 2092

QY 794 TGTCTTGAACCAAGTGTGCTGCTTCAATGTGTTGCTGCTGAACGCGAGATTTGG 853

Db 2091 TGTTTTAACCGTATGCTGCTGAGAACATTTACTGCTGTAGCAGTTTAAAGCCCGAGATTTGG 2032

QY 854 TGAATGCTGTAAGCTATGCTTGAAGACATTTGCTATCTTGAAGGTGATGATGATAC 913

Db 2031 TGAATGACGTAAGCAATGTTAGAACACCAATTAATTAAGTGTGCTCAATGATAC 1972

QY 914 AGAGATCTAGACCTGATTAATAAGTGTACATGACAGCCCTTGACAGGCTGCTAA 973

Db 1971 TGAATGATTAAGTGTGATTAAGTGTGATGCTGATGCTGATGCTGATTAATTA 1912

QY 974 GATTAAGTGTGATTAAGTGTGATGATGCTGATGCTGATGCTGATGCTGATTAATTA 1033

Db 1911 AGTTGAAGTGTGATTAAGTGTGATGATGCTGATGCTGATGCTGATTAATTAAT 1852

QY 1034 TGTCAACGCTATGCTGATTAATGATGATTAATGATTAATGATTAATGATTAATGAT 1093

Db 1851 TGAATGCTGCTGATGCTGATTAATGATTAATGATTAATGATTAATGATTAATGAT 1792

QY 1094 TGAATACTTGAAGACGTTTGGCGAATTTAGTGTGCTGCTGCTGCTGCTGCTGCTG 1153

Db 1791 AGAAATAATTAAGACGATTTGGCAAACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1732

QY 1154 AGCTCCACAGACAGCTTAAATAAAGTGAATGAACTTTCGATGAGATGCTGCTGCTGCT 1213

Db 1731 GGTGCTGATGATTAAGACGCTTAAAGACGCTTAAAGACGCTTAAAGACGCTTAAATTC 1672

QY 1214 TACACGTGACGCTTGAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1273

Db 1671 AACAGCTGGCGGCTGGAAGAGTATCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1612

QY 1274 TATGAAAAGTGAAGCTTGTGAGCTTGAAGGCGATGATGCTACTGACGTAATCTGT 1333

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2001, 19:54:40 ; Search time 97.06 Seconds

(without alignments)
6428.763 Million cell updates/sec

Title: US-09-001-737-7

Perfect score: 1661
Sequence: 1 GAATTCGGCTTCATATGCA.....TGGCGGATTAACCGCAATTC 1661

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 480022 seqs, 187831343 residues

Total number of hits satisfying chosen parameters: 960044

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: N_Geneseq_36.*
2: /cgn2_2/gcgdata/geneseq/geneseqn/NA1980.DAT.*
3: /cgn2_2/gcgdata/geneseq/geneseqn/NA1981.DAT.*
4: /cgn2_2/gcgdata/geneseq/geneseqn/NA1982.DAT.*
5: /cgn2_2/gcgdata/geneseq/geneseqn/NA1983.DAT.*
6: /cgn2_2/gcgdata/geneseq/geneseqn/NA1984.DAT.*
7: /cgn2_2/gcgdata/geneseq/geneseqn/NA1985.DAT.*
8: /cgn2_2/gcgdata/geneseq/geneseqn/NA1986.DAT.*
9: /cgn2_2/gcgdata/geneseq/geneseqn/NA1987.DAT.*
10: /cgn2_2/gcgdata/geneseq/geneseqn/NA1988.DAT.*
11: /cgn2_2/gcgdata/geneseq/geneseqn/NA1989.DAT.*
12: /cgn2_2/gcgdata/geneseq/geneseqn/NA1990.DAT.*
13: /cgn2_2/gcgdata/geneseq/geneseqn/NA1991.DAT.*
14: /cgn2_2/gcgdata/geneseq/geneseqn/NA1992.DAT.*
15: /cgn2_2/gcgdata/geneseq/geneseqn/NA1993.DAT.*
16: /cgn2_2/gcgdata/geneseq/geneseqn/NA1994.DAT.*
17: /cgn2_2/gcgdata/geneseq/geneseqn/NA1995.DAT.*
18: /cgn2_2/gcgdata/geneseq/geneseqn/NA1996.DAT.*
19: /cgn2_2/gcgdata/geneseq/geneseqn/NA1997.DAT.*
20: /cgn2_2/gcgdata/geneseq/geneseqn/NA1998.DAT.*
21: /cgn2_2/gcgdata/geneseq/geneseqn/NA1999.DAT.*
22: /cgn2_2/gcgdata/geneseq/geneseqn/NA2000.DAT.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1661	100.0	1661	20	X86155
2	1052	63.3	1654	20	X86153
3	1035.4	62.3	5365	19	V52210
4	942.6	56.7	3625	20	X12979
5	704.2	42.4	1647	18	T69201
6	704.2	42.4	4972	18	T69203
7	590.6	35.6	1665	20	X86152
8	568.2	34.2	1662	20	X86154
9	552.8	33.3	1635	21	X254509
10	551.6	33.2	910715	20	X20248
11	550	33.1	2465	12	Q13136
12	536.8	32.3	1633	21	A13000

13	536.8	32.3	1635	21	254508	Neisseria meningit
14	528.6	31.8	1838	14	Q47926	hsp60 DNA, Helico
15	528.6	31.8	1838	14	Q48734	Hsp gene, Helicob
16	528.6	31.7	2223	12	O13137	Hyp operon, Chlam
17	527	31.7	1724	19	X14495	H. pylori GHPO 118
18	526.6	31.7	1635	21	254507	Neisseria gonorrhe
19	505.8	30.5	1038602	20	Z01425	Complete genome se
20	496.6	29.9	2284	16	Q90181	Helicobacter pylori
21	496.6	29.9	2284	17	T45681	H. pylori heat sho
22	496.6	29.9	2322	16	O75321	Heat shock protein
23	483.2	29.1	3613	9	N80339	Clone J3178 Insert
24	470	28.3	1017	18	V74577	Staphylococcus aur
25	458.6	27.6	4260	9	N81768	Sequence encoding
26	458.6	27.6	4380	9	N80222	Sequence of Mycob
27	458.6	27.6	4380	19	V05708	Mycobacterium tube
28	446.8	26.8	1626	20	Z11371	Nucleotide sequenc
29	445.8	26.8	1647	17	T14265	Brevibacterium fla
30	444.8	26.8	1569	19	V34608	M. vaccae antigen
31	444.8	26.8	1569	20	Z11343	Nucleotide sequenc
32	441.4	26.6	580073	18	T58840	Mycoplasma genital
33	432.6	26.0	2193	11	O04669	Heat shock protein
34	420.6	25.3	2668	13	O22485	groEL-1 gene. Str
35	419.8	25.3	1620	13	O22482	groEL-1 gene codin
36	409.6	24.7	2242	18	T58403	Human heat shock p
37	362.8	21.8	2167	13	O22484	groES el operon w
38	354	21.3	1320	13	O22481	Heat shock protein
39	315.8	19.0	1777	19	V59425	Heat shock protein
40	296.4	17.8	831	20	V90856	Nucleotide sequenc
41	273.8	16.5	985	20	Z11372	Nucleotide sequenc
42	271.8	16.4	927	19	V34610	M. vaccae antigen
43	271.8	16.4	927	20	Z11345	Nucleotide sequenc
44	217.4	13.1	692	20	X20722	Polynucleotide seq
45	198.8	12.0	544	18	V75108	Staphylococcus aur

ALIGNMENTS

RESULT 1		
ID	X86155	standard; DNA: 1661 BP.
XX		
AC	X86155;	
XX		
DT	22-SEP-1999	(first entry)
XX		
DE	DNA encoding a Streptococcus pyogenes heat shock protein (Hsp)60-2.	
XX		
KW	Heat shock protein; Hsp60-2; immune response; immunological carrier;	
KW	cancer control; tumour; sarcoma; cancer; gene therapy; ss.	
XX		
OS	Streptococcus pyogenes.	
XX		
PN	WO935270-A1.	
XX		
PD	15-JUL-1999.	
XX		
PE	29-DEC-1998;	98MO-CA01203.
XX		
PR	31-DEC-1997;	97US-0001737.
XX		
PA	(STRE-) STRESSGEN BIOTECHNOLOGIES CORP.	
XX		
PI	Mizzen L, Wisniewski J;	
XX		
DR	WPI: 1999-430397/36.	
DR	P-PSDB; Y23904.	
XX		
PT	New nucleic acid encoding heat shock protein-60 from Streptococcus,	
PT	useful in vaccines, as carriers for other immunogens, as anticancer	
PT	agents and for diagnosis	
XX		
PS	Claim 3; Fig 4A-B; 176pp; English.	

XX The present sequence encodes a heat shock protein, designated Hsp60-2.
 CC The protein, its fragments, variants and fusion proteins, are
 CC used to elicit or enhance an immune response against Streptococcus,
 CC and to elicit a similar response to a target antigen fused to the
 CC protein. Unlike other immunological carriers, Hsp60 proteins are not
 CC immunosuppressive so provide an increased response to any conjugated or
 CC fused antigen. Also, where used for cancer control, they lack the side
 CC effects associated with endotoxins. They can also be used to detect
 CC specific antibodies and in treatment or prevention of tumours
 CC (e.g. sarcoma or cancers of breast, ovary, prostate, lung, pancreas or
 CC liver). The Hsp60 polynucleotide is used for recombinant production
 CC of the protein, as a source of primers and probes for detecting
 CC streptococci in standard hybridization/amplification assays, and
 CC therapeutically in gene therapy vectors.

SQ Sequence 1661 BP; 512 A; 308 C; 407 G; 434 T; 0 other;

Query Match 100.0%; Score 1661; DB 20; Length 1661;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1661; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCGGCTCATATGGCAAAAGAAATCAATTTTCAGAGATGGCGCTGCGCATGG 60
 DB 1 gaattcggctcatatggcaaaagaaatcaattttcagagatggcgctgcatgg 60
 QY TGGCGGAGTGTATGATAGCAGATACCGTCAAGTAACGCTTGGCTCAAGAGGCGCA 120
 DB tggcggagtgtatgatatgacagataccgtcaagtaacgcttggctcaagaggcgca 120
 QY 61 tggcgagtgatgtatgtatgtacgacgataccgtaagtaagcttgctccaaaggcgca 120
 DB 61 tggcgagtgatgtatgtatgtacgacgataccgtaagtaagcttgctccaaaggcgca 120
 QY 121 ATGTTTCTCTGAAAAAGCTTTTGGTTTCCCTTAATTAATGACGGGGTACCATTTG 180
 DB 121 atgttctctgaaaaagcttttggtttcccttaatttaattgacggggtaaccattg 180
 QY 181 CTAAAGAGATCGAATTAAGATCATTTTGAAGAACATGGAGCAAAATTCGTCGAAG 240
 DB 181 ctaaagagatcgaaatgaagatcattttgaagaacatggagcaaaatttcgtcgaag 240
 QY 241 TGGCTTCTAAACCAATGATATTGCTGTGATGGAGGAGTACTGCAACAGTTTGAAC 300
 DB 241 tggcttctaaaccaatgatattgctgtgatggaggtactgcaaacagtttgaac 300
 QY 301 AAGCCATTGTTCTGTAAGACCTAAAAATGTGCACAGCGTGCTTAATCCATTGGTATCC 360
 DB 301 aagccattgttctgtaagacctaaaaatgtgcacagcgctgtaattccattggatcc 360
 QY 361 GTCGAGGCAATGTAACAGCAACAGCAAGCTTTGAAGCCTTGAAGCCATTGCTCAAC 420
 DB 361 gtcgaggcaatgttaacagcaaacagcaagctttgaagccttgaagccattgctcaac 420
 QY 421 CTGTATCTGGCAAGAACTATTGCTCAGGTGCGTCAAGTATCATCAAGCTTGAAGAA 480
 DB 421 ctgtatctggcaagaaactattgctcaggtgctgcaagtatcatcaagccttgaagaa 480
 QY 481 TTGAGAGATATATCTAGAAGCTATGAGCGTGTGGCAACGATGGTGATTAACATCG 540
 DB 481 ttgagagatatatctagaagctatgagcggtgtggcaacgatggtgatttaacatcg 540
 QY 541 AAGAACTCTGAGGTATGAAACGAACTTGAAGTGTGAAGCAGCAATTTGACCGTG 600
 DB 541 aagaaactctgaggtatgaaacgaaacttgaagtgtgaagcagcaatttgaccgtg 600
 QY 601 GTTACCTGTCGAATACAGGTGCACAGACATGAAGAAATGTTGCGAGCTTGAAGAAC 660
 DB 601 gttacctgtcgaatacaggtgcacagacatgaagaaatgttgcgagcttgaagaac 660
 QY 661 CATTTATCTTAAATACAGGATTAAGAAAGTGTCAACATCCAGACATTTTGCCACTATTG 720
 DB 661 catttatctttaaatacaggatgaagaaagtgtcaaacatccagacattttgccactattg 720
 QY 721 AGGAAGTTCTTAAACCAACCGTCCATTTACTATTTCAGATGATGTGGATGGTGAAG 780
 DB 721 aggaagttctttaaaccacacgctccatcttactcatltagcagatgatgtgagtgaaag 780

DB 721 aggaagttctttaaaccacacgctccatcttactcatltagcagatgatgtgagtgaaag 780
 QY 781 CACTTCACACCTTGTCTTGAACAGATTCGATGTTCAATGTGTGCTGTCAAG 840
 DB cacttcacaccttgtcttgaacagattcgatgttcaatgtgtgctgtcaag 840
 QY 841 CGCCAGAGATTTGGTATCGTCGTAAGCTTATGCTTAAGACATTTGCTTACACAGTG 900
 DB 841 cgccagagatttggtatcgtcgttaagcttattgcttaagacatttgcttaccagtg 900
 QY 901 GTACAGGATTTACAGAGATCTAGGACTTAATTAAGTATGCTACATACAGCCCTTG 960
 DB 901 gtacaggttttacagagatctaggacttaattagctacatacagcccttg 960
 QY 961 GACAGCGTCTAAGATTACAGTTGATTAAGATAGCACAGTAATTTGTAAGTTACAGAA 1020
 DB 961 gacagcgctctaagattacagttgatataagatagcacagtaatttgtaagttacagaa 1020
 QY 1021 GTTCAGAGCTATTGCTTAACCGTATTGCACTGATTAATGCAATTGAAGAACAACTT 1080
 DB 1021 gttcagagctattgcttaaccgtattgcaactgattaatgcaattgaagaacaaactt 1080
 QY 1081 CTGACTTTGACGCTGAAAAAACTAAGAAAGCTTTGGGAAATTAAGCTGCTGTAGCTG 1140
 DB 1081 ctgactttgacgctgaaaaaaactaagaaagctttgggaaatataagctgtgtagctg 1140
 QY 1141 TTATCAAGATGAGAGCTCCACACAGACAGACTTAAAGAAATGAAACTTCGATTAGAG 1200
 DB 1141 ttatcaagatgagagctccacacagacagactttaaagaaatgaaacttcgattagag 1200
 QY 1201 ATGCTCTAAATGCTACACGTGCAGCCGTTGAGAAGCTATGTTGCTGTGTGGAACAG 1260
 DB 1201 atgctctaaatgctacacgtgcagccgttgagaagctatgttgcgtgtggaacag 1260
 QY 1261 CACTTTTACGTTTATGAAAAAGTAGACAGCTTACAGCTTGAGGGGAGATGATGATCTG 1320
 DB 1261 cacttttacgtttatgaaaaagtagacagcttacagcttgagggagatgatgctg 1320
 QY 1321 GACGTAACTTGTGCTGCTCTAGAAAGAGCGCTGTACGTCAAAATTTGCTTAATGCTG 1380
 DB 1321 gacgtaaactgtgctgctctagaaagagcgctgtacgtcaaaatttgccttaattgctg 1380
 QY 1381 GGTACAGAGGCTCCGTATTATTGACAGTTGAAAAACAGCCCTGCAGAACAGATTTA 1440
 DB 1381 ggtacagaggctccgtattattgacagttgaaaaacagccctgcagaaacagatttta 1440
 QY 1441 ATGCTCAACAGAGTGTGATGATGATTAAGACAGAAATCATTTGACCCGTCAAG 1500
 DB 1441 atgctcaacagagtgtgattgattgataagacagaaatcatttgacccgtcaag 1500
 QY 1501 TAACAGATCAGCGCTTCAAAATGACAGCTTCTGTAGCTAGTCTTATTGACAACAGAG 1560
 DB 1501 taacagatcagcgcttcaaaatgacagcttctgtagctagcttatttgacaacagag 1560
 QY 1561 CAGTTTGTGCTAATTAACCTGAACGCTACGACGCGCCAGCAATGCCAGGATATGG 1620
 DB 1561 cagtttgtgctaattaacctgaacgctacgacgcgccagcaatgccaggatattgg 1620
 QY 1621 ATCCAGGAATGATGGTGGATGGGCGGATTAAGCCGAATTC 1661
 DB 1621 atccaggaatgatgggtggatgggcggttaagccgaatttc 1661

RESULT 2
 X68153
 ID X86153 standard; DNA; 1654 BP.
 XX X86153;
 AC 22-SEP-1999 (first entry)
 DT
 XX
 DE DNA encoding a Streptococcus pneumoniae heat shock protein (Hsp)60-2.
 XX

D	b	1441	aagcgcgaactgaggcgatgggttgataacatgatcatgaagtatcatcatgccagttcaag	1500
O	y	1501	TAAACGATCAGCGCTTCAAAATGCAGCTTGTAGCTACTATTATTGGACAACAAG	1560
D	b	1501		1560
O	y	1561	CAGTTTGTTCTAATAAACCTGAACCAGCTACGCCAGCCGACAGAATGCCAGCATGTGC	1620
D	b	1561		1620
O	y	1621	AATCAAGGATGATGGG	1636
D	b	1621		1636
			tgggtcggaattgggcgg	1636
R	E	S	RESULT	3
I	D	V	52210	
X	X		V52210 standard; DNA; 5365 BP.	
A	C		V52210;	
D	T		23-OCT-1998 (first entry)	
X	X		Streptococcus pneumoniae genome fragment SEQ ID NO:77.	
K	M		Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;	
X	X		computer readable medium; vaccine; pharmaceutical composition; ds.	
O	S		Streptococcus pneumoniae.	
P	N		WO9818931-A2.	
P	D		07-MAY-1998.	
P	F		30-OCT-1997; 97MO-US19588.	
P	R		31-OCT-1996; 96US-0029960.	
P	A		(HUMA-) HUMAN GENOME SCI INC.	
B	L		Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M,	
K	S		kunsch CA, Rosen CA,	
W	P		WI; 1998-272225/24.	
C	C		Computer-readable medium with recorded Streptococcus pneumoniae	
P	T		polynucleotide sequences - useful in diagnostic kits and assays, and	
P	T		pharmaceutical compositions and vaccines for Streptococcus	
P	T		pneumoniae	
X	X		Claim 1; Page 628-631; 1409pp; English.	
X	X		The present invention describes a computer readable medium which has	
C	C		the nucleotide sequences SEQ ID NO:1 to 391 (V52134 to V52524) recorded	
C	C		on it, or a representative fragment or a sequence at least 95% identical	
C	C		to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in SEQ ID NO:1	
C	C		to 391 (V52134 to V52524) are genomic fragments from Streptococcus	
C	C		pneumoniae. The present invention also describes an isolated nucleic acid	
C	C		molecule encoding a homologue of any of the fragments of the S.pneumoniae	
C	C		genome (SEQ ID NO:1 to 391) where the nucleic acid molecule is produced	
C	C		by a process comprising: (a) screening a genomic DNA library using as a	
C	C		probe a target sequence defined by any of the sequences in SEQ ID NO:1	
C	C		to 391, identifying members of the library which contain sequences	
C	C		that hybridise to the target sequence and isolating the nucleic acid	
C	C		molecules from the members; or (b) isolating mRNA, DNA or cDNA produced	
C	C		from an organism, amplifying nucleic acid molecules whose nucleotide	
C	C		sequence is homologous to amplification primers derived from the	
C	C		fragment of the S. pneumoniae genome to prime the amplification and	
C	C		isolating the amplified sequences. The computer readable medium can be	
C	C		used in a computer-based system for identifying fragments of the	
C	C		S. pneumoniae genome of commercial importance, or expression modulating	
C	C		fragments of the S. pneumoniae genome. Products from the present	
C	C		invention can be used in diagnostic kits and assays, and pharmaceutical	

[illegible]

QY	974	GATTACAGTTGATTAAGATATGCACACATTAATTTGGTGAAGTTGACGAGTCAAGATAT	1033
Db	1237	agtagccgtyggacaaagaatagacccgctcatctgtagaaggctgacgaaatccctgaagcgt	1286
QY	1034	TGCTAACCCGATTTTGCACCTGATTTAAATCGCAATTAGAAAACAACACTTCTGCATTGGACG	10933
Db	1297	ttctccaccgctggttcggtttatcaagctccaatcgaaactataactctctgaatttgaccg	1356
QY	1094	TGAAAAACTACAAAGACGTTTGGCCGAAATTTAGCTGGTGGTGTAGCTGTTATTCAAAGTAGG	11533
Db	1357	tgaanaattgaaagaacgcttggccaattctgaattgcaagtgtgtagcggttattaagtttgg	1416
QY	1154	AGCTCCACAGACGACGACTTTAAATAAATAATGAAACTTCGCATTGGAGATGCTCTAAATGC	12133
Db	1417	agcccgcaactgaaactctgagtttgtaaagaataatgaaactccgcacttgtaagaatgtccctcaacgc	1476
QY	1214	TACACGTCAGACGCCCTTAAACAAGTATTCGTCGTGCTGTGCAACAGCACTTATTACGT	12733
Db	1477	taccctgycagcttgcttgaagaagttatctgttgcaggtggtggaacagctcttggccaatgt	1536
QY	1274	TATTGAAAAAGTACACACTCTCTTGAAGCTTGAGGCGATGATGCTACTGACGTAACATTGT	13333
Db	1537	gattccagctggttgcctactaccttgaattgaacgagatgtagaacaacgaagcaatattgtc	1596
QY	1334	GCTTCGTGCTCTCAGAGAGCGCTGTACGTCATAATGCTTTAAATGCTGGGTACGAAGCTC	13933
Db	1597	tctccgtgcttcttgaagaaccgcgtctgtccaattgtccacaattgcagaagatttgaagatc	1656
QY	1394	CGTAGTTATTGACACAGTTTGAATAAACAGCCCTGACGAAACAGATTTAATGCTGCAACAGG	14533
Db	1657	tatcgttatcgaatcgtcttgaaaaaatgcctgaacttggatlaagattttaaocgaacgaactcg	1716
QY	1454	TGAGTGGGTTGATATGATTTAAACAGGAAATCATTTGACCTGTCAACATTAACACGATCAGC	15133
Db	1717	cgaatgggttaacatgattatgataagatcatcatgataccagttaaagtgagctgtcagc	1776
QY	1514	GCTTCAAAATGACAGCTTCTGTAGCTAGCTTATTTTGACACAGAGAGATTGTTGCTAA	15733
Db	1777	cctcaaaaatgacgacatctgtagccagcttgattttgacaacagaagcagctgtagccaa	1836
QY	1574	TAAACCTGAAACACAGCTACGGCCACGCGCAGCAATGCGCAGAGGTATGATCCAGGAATGAT	16333
Db	1837	taaacacgaacaccagtagccccaactccagaatgataccaaacatcatgatgagcgagatgat	1896
QY	1634	g 1634	
Db	1897	g 1897	
RESULT 4			
X12979			
ID X12979 standard; DNA; 3625 BP.			
X12979;			
19-MAR-1999 (first entry)			
XX	DE	Enterococcus faecalis genome contig SEQ ID NO:42.	
XX	XX	Enterococcus faecalis; contig; detection; Enterococcal infection;	
KW	RV	vaccine; attenuation; computer readable medium; de.	
XX	OS	Enterococcus faecalis.	
XX	XX	W09850555-A2.	
XX	PN	12-NOV-1998.	
XX	PD	04-MAY-1998; 98MO-US08985.	
XX	PF	14-NOV-1997; 97US-006609.	
RR	RR	06-MAY-1997; 97US-004431.	
RR	RR	16-MAY-1997; 97US-004655.	

P	A	(HUMA-)	HUMAN GENOME SCI INC.
X	X		
P	I	Barash SC,	Dillon PJ, Kunsch CA;
X	X		
D	R	WPI:	1999-045171/04.
X	X		
P	T	New isolated Enterococcus faecalis polynucleotides and polypeptides	
P	T	- used to develop products for the detection of Enterococcus and for	
P	T	use in vaccines for prevention or attenuation of Enterococcus	
P	T	infection.	
P	S	Claim 1:	Page 414-416; 2084pp: English.
X	X		
C	C	A computer readable medium has been developed which has recorded on it	
C	C	982 nucleotide sequences isolated from the Enterococcus faecalis genome.	
C	C	XI2938 to XI3919 represent these nucleotide sequences which are primary	
C	C	nucleotide sequences, also known as contigs. The computer-based system	
C	C	can identify fragments of the Enterococcus faecalis genome with	
C	C	commercial importance. The products can be used to detect the presence	
C	C	of Enterococcus faecalis in samples. They can also be used for	
C	C	diagnosing Enterococcal infection in an animal and monitoring	
C	C	progression of disease, and for identifying agents which can be used to	
C	C	modulate the growth or pathogenicity of Enterococcus faecalis, or	
C	C	another related organism, in vivo or in vitro. In particular the	
C	C	polypeptides encoded by the Enterococcus faecalis nucleotide sequences	
C	C	can be used in vaccines to prevent or attenuate an Enterococcal	
C	C	infection.	
X	X		
S	Q	Sequence	3625 BP; 1335 A; 542 C; 750 G; 993 T; 5 other;
		Query Match	56.7%; Score 942.6; DB 20; Length 3625;
		Best Local Similarity	74.2%; Pred No. 4.7e-237;
		Matches 1191; Conservative	0; Mismatches 415; Indels 0; Gaps 0;
O	Y	15 ATGGCAAAAGAAATCAAATTTCACAGATGCGCTGTGCCATGTGCAGCGAATGAT	74
D	b	364 atggscaaaagantlaaatlgttcagaatgaatgcctcgaacgatcgtagtgat	423
.O	Y	75 ATGTTGACGAGTACCCTGAACACTAACCGTTGTCTCTTAAGGGCCGCAATGTTCTTGAA	134
D	b	424 gatltagcagatacagltgaagtgcattagccctaaggcslgaactgclttlagaa	483
O	Y	135 AAAGCTTTGGTTCGCCCTATTAATACTAAATGCGGGGTAACCATTTGTAAGAAGATCCAA	194
D	b	484 aaatcatltyglaccacctlgtacttaacgaatgagtaaacaatlyccaagyaatltaa	543
O	Y	195 TTAGAAGATCATTTTGAAAACAATGGAGCAAAATTTGTGTGGAAGTGGCTTTAAACC	254
D	b	544 ttgggaagatcatlittgaaaacaatgyggycacaatatagtlitccagaagtlcgcttcaaacy	603
O	Y	255 AATGATATTTGCTGTGGATGGAGCAGCTACTGCAACAGTTTTGACACAAGCCATTGTTGAT	314
D	b	604 aatgatattgctgtgacggaaccaacaacagactglttligacaagaacatlgctgt	663
O	Y	315 GAAGGACAAAAAATCTGACACACAGCTGATATCCAATTTGGATTCCTCGAGGACTTGAA	374
D	b	664 gaaggtctaaaaaacgttaactgtcygaacccaacatlagtatctgcgtgysgtllgae	723
O	Y	375 ACAGAACACGACGACGCTGTTGAACCTTTGAAGCAATTTGCTCACACCTGTATCTGGCAAG	434
D	b	724 ttagcaacaaaacacgcgttagaagaatatacaacaatalttcatctgttgttgcattaaba	783
O	Y	435 GAAGCTATTGCTCAGTGCCTGCTGACATATCACCCTCTGAAAAGATTGAGAGATATAT	494
D	b	784 gaagcattgcacaacgctgcgtcgtlctalcacagtlcttgaanaagctgcgccaatlaatt	843
O	Y	495 TCAGAAAGCTATGAGCGTGTGGGCAACGATGGTGTGATTTACATGGAAGATCTCGAGCT	554
D	b	844 gccagctlgaacatlyttaaaaatlytgtaacagcgctaaatcacattytgaatccaaaayg	903
O	Y	555 ATGGAAACGCACTTGAAGTGGTTGAAGGACATGTGACCTGTGGTGTACCTGTCTGAA	614

```

Db 904 attgaacagaatcattagatggtctgaagaaatgcgaatccgcgggtattatctca 963
QY 615 TACATGGTCACAGACATGAATAAATGGTTGCAGACCTTGAATAACCATTTATCTTAATC 674
Db 964 taactggtactgcaacagataaaatggaagcgtctttagaataatccatatacttaatt 1023
QY 675 ACGGATAAAAAGTGTCAACATCCAGACATTTTGGCAGTCTTGGAGGAATCTTAA 734
Db 1024 accgacaaataatctcaaatatcaagatatcttaaccttattagaaataatctcaaa 1083
QY 735 ACCAAGCGTCATTAACATTAATTCAGATGATGATGATGATGATGATGATGATGATGAT 794
Db 1084 caaagccgtccacattgattatcgatgagatgagatgagatgagatgagatgagatgagat 1143
QY 795 GCTTTGAACAAGATTCGTGCTACTTCTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 854
Db 1144 gtattgaaacaaatccggtgacatttaattgttcgcgaagtaaaagccgcaagatttgg 1203
QY 855 GATGCTGTAAGCTATGCTTGAAGACATTCATCTTGGACAGTGTGTAAGTATTAACA 914
Db 1204 gaccgcgcgaagcagatctgaaagatttgcattttaaacaggtggtacagtaact 1263
QY 915 GAGATCTAGACCTTGAATTAAGATGCTACATGACAGCCCTTGGACAGCTGCTAG 974
Db 1264 gacgacttaggttagagattaaagaacaacattgaaacttaggaatgctacga 1323
QY 975 ATTACAGTTGATTAAGATGACAGATTAATTTGATGAGTGTGAGTGTGAGTGTGAGTGTGAG 1034
Db 1324 gtatgttcgcgaagataaacaacatttgcgaaggtgctgttcataaagaagccatt 1383
QY 1035 GCTAACCGTATTTGACATGATTAATTCGCAATTAACAACACACTTCTGACTTGAACCGT 1094
Db 1384 gatcccgcttcttaatttaaaacaaatcgcaaaacagctcgtatttgcatt 1443
QY 1095 GAAATACTACAGAGAGCTTTGGCAATTTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1154
Db 1444 gaaataatcaagaacgttctgcaaatctagctggtggtgctgctgctgctgctgctgctgct 1503
QY 1155 GCTTCAACAGAGAGAGCTTTAAAAAATGAACCTTGCATTTGAGATGCTCTTAATGCT 1214
Db 1504 gctgcaactgaaacagataaaagaattaaagaattcaagaattgagatgcaataacgca 1563
QY 1215 ACAGCTGACAGCCCTTGAAGAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1274
Db 1564 acacggtccgcttagaagaagagcattgctggtggtggtggtggtggtggtggtggtggt 1623
QY 1275 ATTGAAAGATGACAGCTTTGAGCTTGAAGGCGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1334
Db 1624 attggttaagtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1683
QY 1335 CTTCGCTCTAGAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1394
Db 1684 gtctgctcattagaagaacaaatccgtcgaatcgtcgaatcgtcgtcgtcgtcgtcgtcgtcgt 1743
QY 1395 GTACTTATTGACAACTTGAAGAAGAGCCCTGACGAGAACAGATTTAATGCTGCAACAGGT 1454
Db 1744 gtgattgttgaacaaataaataatgctgactagatgctgactgactgactgactgactgact 1803
QY 1455 GAGTGGTGTATATGATTAATAACAGATCATTTGACCTGTGCAAGTAAACAGATCATGAGC 1514
Db 1804 gaaatggttaacacatgcttgaagccgcatgctgacccaaacaaagaatcgtctgctgctgct 1863
QY 1515 CTTCGAATGACAGCTTCTGAGTACTGATTTATTTTGAACAACAGAGATGCTGCTGCTGCTGCT 1574
Db 1864 ttcaacaaatgagcttctgctgctgcttattatcaacaaatgagaaatgctgctgctgctgct 1923
QY 1575 AAACCTGACACGCTAGCCAGCCAGCAATGCCAGCAGATGAG 1620
Db 1924 aaacacgaacacagctgacacagctcctcatgactgactcactcaatgag 1969

```

RESULT 5

```

T69201
ID T69201 standard; DNA; 1647 BP.
XX
AC T69201;
XX
DT 20-AUG-1997 (first entry)
DE
DE Lawsonia intracellularis GroEL DNA.
XX
KW Intestinal disease; porcine proliferative enteropathy; vaccine;
KW GroEL; heat shock protein; ss.
XX
OS Lawsonia intracellularis.
XX
PN W09720050-A1.
XX
PD 05-JUN-1997.
XX
PE 29-NOV-1996; 96WO-A000767.
XX
PR 30-NOV-1995; 95AU-0006911.
PR 30-NOV-1995; 95AU-0006910.
XX
PA (DARA-) DARATECH PTY LTD.
PA (PIGR-) PIG RES & DEV CORP.
XX
PI Hasse D, Panaccio M;
XX
DR WPI: 1997-310605/28.
DR P-PSDB: W16678.
XX
PT Vaccine for treating or preventing Lawsonia intracellularis
PT infection - especially in pigs, containing non-pathogenic form of
PT bacterium or its components
XX
PS Claim 12; Page 38-42; 94pp; English.
XX
CC A DNA molecule (T69201) codes for the GroEL heat shock protein
CC (W16678) of Lawsonia intracellularis, the causative agent of porcine
CC proliferative enteropathy (PPE). A genomic library was prepd. from
CC L. intracellularis PPE lesion isolates and screened with rabbit
CC anti-L. intracellularis antiserum. Phagemid DNA was isolated from
CC individual clones and sequenced. GroEL and GroES (see also T69202)
CC sequences were identified and another 13 clones were subsequently
CC (T69203-15) isolated. These nucleic acids can be used to produce
CC recombinant polypeptides useful in vaccines against intestinal
CC diseases such as PPE, including recombinant vaccines utilizing
CC bacterial, fungal or viral vectors. They can also be used as
CC genetic vaccines and in diagnostic assays.
XX
SQ Sequence 1647 BP; 536 A; 277 C; 368 G; 466 T; 0 other;
XX
Query Match 42.4%; Score 704.2; DB 18; Length 1647;
Best Local Similarity 66.4%; Pred. No. 8; 2e-175;
Matches 1044; Conservative 0; Mismatches 523; Indels 6; Gaps 2;
QY 19 CAAGAAGAACCAATTTTACAGAGATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 78
Db 8 ctaagaacaccccttctgactgacaaagccgtgaaacattcaagagtgatgaataac 67
QY 79 TACGAGATACCGTCAAGTAAAGCTTGGCTTAAGGCGCAATGTTGTTGTAAGAAAG 138
Db 68 ttgcaaatgctgttaagaataacacttgaccataaagccgtaatgctgctgctgctgctgctgctgct 127
QY 139 CTTTGGTCTTCCCTTAATTAATGACGGGTAACCAATTGCTTAAGAGATGGAATTAAG 198
Db 128 ctttggctcccaagatattacaagaatggtgctatctgtgcaaaaagaattgaaacttg 187
QY 199 AAGATCAATTTGAAGAACATGGAGCAAAATTTGCTGCTGTAAGGCGCTTCAAAACCAATG 258
Db 188 aagataagttgaaataatgagcgctcaaatggttlaaagaagtagctcccaaaactagcg 247

```

OY	259	ATHTTGGGTGGATGGGACACACTACTGCAACAGCTTTTGACACAAAGCACTTTGTCTATGAAG	318
Db	248	atactgcctggtgtagaaggacatacaacagcaagctccctgcgaagaactatattacgttcgtgaag	307
OY	319	GACATAAAAAGTGTACAGCGAGCGTGTATTCATTTGGTTCGTGTGAGGCACTTGAAGAACAG	378
Db	308	gtgtaaacctgttagcaagcgtggtccgaataccctatgacctaaacgttgcataagataag	367
OY	379	CAACAGCAACAGCGTGTGAAGCCCTTGAAGCCATTCCTCAACCTGTATNTGGCAAGAGAG	438
Db	368	ctgtgtgtcgtcttaccataaagaactaagacgtacatacaagctactcgttgcacaaag	427
OY	439	CTATTGCTCAGGCTGCTGCAGTATCATACGCTGTGA---AAAAGTTGAGAGTATATCT	495
Db	428	aaatagctcaagttgtgaaccattctgcacaactctatacaacaatagytatcatcatag	487
OY	496	CAGAAAGCATATGGAGCCGTGTGGGACAGCATGSGTGTGATTCACATTCAGAAATCTCAGATGA	555
Db	488	ctgaagcctatgtgctaaagtttggaaaaagagtggtatcacagcttgaggaagctaaagctc	547
OY	556	TGGAGAACGAACTTGAAGTGGTGTGAAGGCAATGCATTTTACCGTGGTTCACGTGTCAAT	615
Db	548	tltgaactaactatgtagtgtgttgaaggaatgtagttagccgttgcctactccctccat	607
OY	616	ACATGCTACAGACATGAATAAATGGTTGCGAGACCTTGAAAAACCATTTATCTTATATCA	675
Db	608	acttgttaactaactcctgagaanaatggtttgtgaacttgatatacccttatcccttgta	667
OY	676	CGGATAAAAAGTGTCAAACATCCAGACACTTTTGGCACTACTTGAAGAGATTCTTAAAA	735
Db	668	atggaataaagatctacagatagaagaacatgctacaaactctagaacaagttgtctaaag	727
OY	736	CCAAACCGCTTTCATCTCATTTATTTGAGATATGTGATGTGATGATGATGATGATGATGATG	795
Db	728	taaacctgcctccctctatattatctgcgtgaagaagctagaaggttgaagcaacttgcacaactg	787
OY	796	TCTTGAACAGATTCGTGTGATCTTTCATGTGTTGCTGTCAAGACGCCAGGATTTGGTG	855
Db	788	tagtcaataagctctcgtggtgagcaactcccaagttgtacgttaaaagctcctgtttgtgtg	847
OY	856	ATCGTCCGTAAGCATATCTGGAAGACATTGTCTATCTTGAACAGTGTGTACAGTGTATACAG	915
Db	848	aacgcgtgaagaagctatgcttgaagaatattgtcatccctactggaaggaagaacatatttg	907
OY	916	AGGATCTAGACCTTGAAATTTAAAGATGCTTCATATGACACACCCCTGGACAGCGTGTCTAGA	975
Db	908	aagatcgtgtatataaagctctgaaaatgttaagcttgcctccctttaggaacaagctaaacgtg	967
OY	976	TTACAGTTGATTAAGATAGCAGATATTTGTAAGGTTTCAGAGAACTTCAGAGCTATTTG	1035
Db	968	tagtatattgacaagaataactactatcgttgaigtgtgtgtaaaatccagaagatata	1027
OY	1036	CTAACCGTATTTGCACTGATTTAAATTCGCAATTTAGAAACAACAACCTTCTGACTTTGACGGTG	1095
Db	1028	aagctcgtagttaaacaacaaattcgtgcacaacatltgaagaacaagctcagaatactgtcgtg	1087
OY	1096	AAAAATCAACAAGAACGTTTGGCGCAAAATTAACGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1155
Db	1088	aaaactctcaagaagcgtcttgcgaacaactctgttggtaggtagctgtatccaactgttggag	1147
OY	1156	CTCCACAGAGACAGCTTTTAAAGAAATGAACATTCCTGCATTTAGAGATGCTCTAAATGCTA	1215
Db	1148	ctgtctactgtaaacctgaatgtgaagagaagagtagctgtgtatagaagaagtctcaaatgtca	1207
OY	1216	CACGTGCAGCCGTTTGAAGAAGGTATGTTGTCGTGCTGTGTGTGTGTGTGTGTGTGTGTGT	1275
Db	1208	caagaagctcgtgttgaagaagtagtctgtccctgtgtgtgtactgtcttcttgcctccca	1267
OY	1276	TTTGAAGAAAGTGTGACGCTCTTGAGCTGTGTGAGCGCATGATG---CTACTGTGACCTATCATTCG	1332
Db	1268	ctaaagctccttgcatactataaacctcgtctgaatgaatgaatgaatgaatgaatgaatgaatga	1327
OY	1333	TGCTTGTGCTCTGTGAAGACCTGTATCGTCAAAATTCCTTAAATGCTGTGGTTCGAAAGCT	1392

Accession	Sequence	Length
Dd	1328 tccgcgcttcctcttgaagagcccttcaacgtatgcycaaatgctgctatgaagct	1387
Oy	1393 CCGTAGTATTATGACACAGCTTGAANAACAGCCCTCAGAGAGAGATTTAATGCTCCAGAC	1452
Dd	1388 ctatgtctgttgaanaaagcttcgtgaaccocaaagatgcttcttgatttaatgctgcatcag	1447
Oy	1453 GTGAGTGGGTGATGATCATTTAAACAGAGATCTTATACCCTGTCAAGATGACAGCATAG	1512
Dd	1448 gagatatagaagacctattataaagctggtgctcttgatctcctaaaaaagttaacgtatg	1507
Oy	1513 CGCTTCAAAATGACAGCTTCTGTAGCTAGTCTTATTTTGCAACAGACAGCATGTTGTCTGA	1572
Dd	1508 catcacaataaagcagcctcagtcagcctccttacttctaactacaataagcgtatttgct	1567
Oy	1573 ATAAACCTGAAC	1585
Dd	1568 aaaaaccagagacc	1580

Result	ID	Accession	Score	DB	Length	Query Match
XX	T69203	T69203 standard; DNA: 4972 BP.	42.4%	DB 18;	Length 4972;	Query Match
XX	AC	T69203;	66.4%	Pred. No. 1.3e-174;		Best Local Similarity
XX	DT	20-AUG-1997 (first entry)	0;	Mismatches 523;	Indels 6;	Gaps 2;
XX	DE	Lawsonia intracellularis vaccine candidate DNA.				
XX	KN	Intestinal disease; porcine proliferative enteropathy; vaccine; ss.				
XX	OS	Lawsonia intracellularis.				
XX	PN	W09720050-A1.				
XX	PD	05-JUN-1997.				
XX	PF	29-NOV-1996; 96WO-AU00767.				
XX	PR	30-NOV-1995; 95AU-0006911.				
XX	PR	30-NOV-1995; 95AU-0006910.				
XX	PA	(DARA-) DARATECH PTY LTD.				
XX	PA	(PICR-) PIC RES & DEV CORP.				
XX	PI	Hasse D, Panaccio M;				
XX	DR	WPI; 1997-310605/28.				
XX	PT	Vaccine for treating or preventing Lawsonia intracellularis				
XX	PT	infection - especially in pigs, containing non-pathogenic form of				
XX	PT	bacterium or its components				
XX	PS	Claim 14; Page 47-50; 94pp; English.				
XX	CC	DNA molecules (T69203-15) were isolated from an immunoscreening of				
XX	CC	a Lawsonia intracellularis library using experimental sera from				
XX	CC	vaccinated pigs. These nucleic acids, as well as isolated GroEL				
XX	CC	and GroES (T69201-02) sequences, encode putative vaccine candidates				
XX	CC	(see also W16678-85) sequences, useful for protection of animals and birds				
XX	CC	against intestinal diseases, esp. protection of pigs against				
XX	CC	porcine proliferative enteropathy (PPE). They can also be used as				
XX	CC	genetic vaccines.				
XX	SEQ	Sequence 4972 BP; 1717 A; 871 C; 831 G; 1532 T; 21 other;				

D	b	2055	ctaaagaatcctcttttcttgatgctaagcccgctgaaaaacttccagaggttgaagtaaac	2114
O	y	79	TAGCAGATACCGGTAAGAAGTAACGCTTGCTGCTTAAGGGCGCAATGTTGTTCTTGA AAAA	138
D	b	2115	ctgcaaatgctgtctaagaatacacttggacctaagaagccgtaaltctgtatctgaaagt	2174
O	y	139	CTTTTGGTTTCCCTTAAATTACTAATGACGGGGTAACTCAATTGCTTAAAGATCGAATTAG	198
D	b	2175	cttttgttccccagttatacaaaaatggtgtatctgttgcgaagaatctgaacttg	2234
O	y	199	AAGATCATTTTGA AAACATGGGACAAATTTGGTGTCTGAAGTGCTTCTTAAACCAATG	258
D	b	2235	aagataagtttgtaaaatalatgagcgctcaaaatggtltaagaagtaagctcccaaaactagc	2294
O	y	239	ATATTGCTGGTGAATGGGACACACTACTGCAACACTTTTGGACACAACCAATTGTTCATGAG	318
D	b	2295	atatgctggtgtagatggaacacacttaacaaagcaacagctccttgcacaagctatltatctgtag	2354
O	y	319	GACTTAAAAATGTATACAGCAGAGTGCTAATTCAAATTGTTCCGTGGAGGCAATTGAAACAG	378
D	b	2355	gtgtaaaactctgtagcaagctggtgtcgtatccctatctgcatctaaacgtgtcagtagtaag	2414
O	y	379	CACACGACACAGCTGTTGAAGCCTTGAAAGCCATTGCTCAACCTGTATCTGCAAGACAG	438
D	b	2415	ctgtgtgtgtcgtgttactaaagaactaagcgacatltacaagaactactcgtgtgacaaaag	2474
O	y	439	CTATTGCTCAGAGTGCGTGACAGTATCATACAGCGCTCTGA---AAAAGTTGAGAGTATTACT	495
D	b	2415	aaatagctcaagttgtagaaccaatctcttgcgaacactctgtatacaaatagtgtaatacatag	2534
O	y	496	CAGAAGCTATGGAGCGTGTGGGCAACGATGCTGTGATTACATCGAAGATCTCGAGGTA	555
D	b	2535	ctgaagctatgtctaaagtttggaaaaggaagtgctatctacagttgagaagctlaaagctc	2594
O	y	556	TGGAAACAGACCTTGAGTGTTGAAGGCACTGAATTAACGCTGGTATACGCTGCAT	615
D	b	2555	ltgaaactaacctagaatggtgttgaagaagatgaaatcttgcgtggtacccctccatc	2654
O	y	616	ACATGCTCACGACAAATGA AAAATGGTTGTCAGACCTTGA AAAACCCATTATCTTAAATCA	675
D	b	2655	acttgttaactaatccctgagaaaatggttltgtgaacttgataaacttatalactcttltgta	2714
O	y	676	CGGATAAAAAGTGTCAAACTCCAGACACTTTTGGCACTACTTAGAGAACTTCTTAAA	735
D	b	2715	atgagaaaaaagatctacatagatagaataagacatgttaccaaacttgaacaagtttctaag	2774
O	y	736	CCAACCGGCCATTTCGATATTATTGACAGATGATGTGATGCTGACACCTCCACACCTTG	795
D	b	2775	taaacgtccctcaactatattatgtctggaagaagtagaagatggaagcactltgcaacacttg	2834
O	y	796	TCTTGAACACAGATCGTGGTACTTTCAATGCTGGTGTCTGCAAAAGCCGACGAAATTGGTG	855
D	b	2835	tagtcaataaagctctcgttggagcaactccaagtgttgaacgtlaaaagctccctgttlttgg	2894
O	y	856	ATCGCGTAAAGCATATGCTTGGAAGACATGCTATCTGTGACAGGTGTACAGATTAACAG	915
D	b	2895	aaagcgcgaagaactatgtcttgaagaatctgtcatcttcaatcgtgaagagaagaaqaataattg	2954
O	y	916	AGGATCTAGGACTTGAAATTA AAAAGATGCTACATGACACCCCTTGAGACGGCTGTCAAGA	975
D	b	2955	aagatcgtgtatgaataagcttgaaaaatgtlaagcttcttctttaggaacagctlaaagctg	3014
O	y	976	TTCACGTTGATAAAGATAGACACAGTAAATTGTTGAAGGTTACAGGAAGTTACAGACTATTG	1035
D	b	3015	tagtatgataaaaagaaaatactacatacgttgaatgtgtgttgcgtgaaaaatcagaagatla	3074
O	y	1036	CTAACCGGATATGCTGATTAATTAATTCGAATTAAGAAACAACACTTCTGACTTTGACCGTG	1095
D	b	3075	aagctcgtgattaaacaatctcgtgcacaaattggaagaaaacaaagcttcagatatgaatcgtg	3134
O	y	1096	AAAACTACAAAGACGTTTGGGCAAAATTAGCTGTGTGTACTCTGTTATCAAAAGTAGAG	1155

Db	3135	aaaaacttcaagaacgctcttgcacaaacttctgttgagtagtctgttcatccatgcttgcag	3139
Qy	1156	CTCCACGAGACACAGCCTTTAAAGAAATGAAACTTGGCATTTGAGATGCTCTTAATGCTA	1215
Db	3195	ctgtctacttgaacctgaactgaaatgaaagagaagaatcgttctgaaagatgtctctaaatgcaa	3255
Qy	1216	CACGTGCAGCGCTTGAGAAAGGATCGTCTGTGCTGTGGTGAACAGCAGCTATTACGGTTA	1275
Db	3255	caagaagcgtcggttctgaaagaagatctgtcccttggctgtgtactgtcttcttccgtcca	3314
Qy	1276	TTGAAAAGATGACGACGTCCTTGAGCTTGAGGGCGATGATG--CTACTGGACGTTAACATTG	1335
Db	3315	ttaaagtcctctgatalgatataataaacctgtcatgatgatgaacttctgtgaacttaataca	3374
Qy	1333	TGCTCTGCGTCCTGAGAGAGCGCTGTACGTCAAATTCCTTTAAATGCTGGGTACGAAGCT	1392
Db	3375	tccgctcgcttccctcttgaagaagccttcttaacgtatcgtcgaacatgctggtctatgaagct	3433
Qy	1393	CCGTGATTATTGACAGATTGMAAAMACAGCCCTCCAGAGACAGGATTTAATGCTGCACAG	1452
Db	3435	ccatitgttctgtaaaaagcttcgttgaaccaaagaatgttcttgatttaattgctgcatacg	3494
Qy	1453	GTTGATGGGTTGATATGATTAAAAACAGGATTCATTGACCTGTCAAGATTAACAGCATCAG	1512
Db	3495	gagaatattgaagacctattataaagcttgglygtcatgtatccctaaaaaagttacaacgtatg	3555
Qy	1513	CGCTTCAAAAATGACAGCTTCTGTGATGCTGTATTATTGACAAACAGACAGATTTGGCTA	1572
Db	3555	cattacaaaatgaacgatactgataagccttcttaacttctaatacaagaatgctgattgtctg	3614
Qy	1573	ATTAACCTGGAAC 1585	
Db	3615	aaaaccagaaacc 3627	
RESULT: 7			
ID	X86152	standard; DNA; 1665 BP.	
XX	XX	X86152;	
AC	XX		
XX	XX		
DT	22-SEP-1999	(first entry)	
DE	XX	DNA encoding a Streptococcus pneumoniae heat shock protein (Hsp)60-1.	
XX	XX		
KM	XX	Heat shock protein; Hsp60-1; immune response; immunological carrier;	
XX	XX	cancer control; tumour; sarcoma; cancer; gene therapy; ss.	
OS	XX	Streptococcus pneumoniae.	
XX	XX		
PN	XX	W09935270-A1.	
PD	XX	15-JUL-1999.	
XX	XX		
PF	29-DEC-1998;	98WC-CA01203.	
XX	XX		
PR	31-DEC-1997;	97US-0001737.	
XX	XX		
STRE-	XX	STRESSGEN BIOTECHNOLOGIES CORP.	
XX	XX		
PI	XX	Mizzen L, Wisniewski J;	
XX	XX		
DR	XX	WPI: 1999-430397/36.	
XX	XX	P-PSDB: Y23901.	
XX	XX		
XX	XX	New nucleic acid encoding heat shock protein-60 from Streptococcus,	
XX	XX	useful in vaccines, as carriers for other immunogens, as anticancer	
XX	XX	agents and for diagnosis	
XX	XX		
XX	XX	Claim 3; Fig 1A-B; 176pp; English.	
CC	XX	The present sequence encodes a heat shock protein, designated Hsp60-1.	
CC	XX	The protein its fragments, variants and fusion proteins, are	

[illegible]

Matches	958:	Conservative	0:	Mismatches	633:	Indels	6:	Gaps	27:
QY	18	GCAAAAGAAATCAATTTTTCAGCAGATGCGTCTGCCATGTGCGCGAGTTGATATG							77
Db	7	gaaaaaagcagtaatttcgcaatgtaagtcgcgcaaaaaaatcgttgaaacgctgtaacatt							66
QY	78	TTAGCAGATACCGGTCAAAAGTAAACGCTTGCTCTAAAGGGCGAATGTGTTCTTGAAAAA							137
Db	67	ctgagcaaaagccgctgcgctaaaccttggtctcccaaaagccgcaacgctggtgtgacgcgc							126
QY	138	GCCTTTGGTTCTCCCTTAATTTACTATATGCGGGGTAACCATGTGCTAAAGAGATGCAATTA							197
Db	127	gcttcgcgggcgcgcacatcaccaagaagcggtacacgcgtcccaaaaaaatcgaaacg							186
QY	198	GAAGATCATTTTGAAGAAACATATGGACCAAAATATGTGTGTGAAGTGGCTTCTTAAACCAAT							257
Db	187	aaagcgaagcttgtaaaatatgtggtgcgcgaatatggtgaagaagctgcgtcccaaaaccc							246
QY	258	GATATTTGCTGGTGTGATGGACGACTCTGCACACAGTTTGTGACAAAGCCATTGTTCAATGA							317
Db	247	gacgtgycgggycgaagctcgcactacgcgcacccgctaltgtgcgaatccatccgttcgcga							306
QY	318	GCACGTAAAAAATGTGACACGACGCTGCTAATCCAAATTTGATATCCGTCGAGGCACTTGACA							377
Db	307	ggtatgaataacgtaacgtaacgcgcgtaltgaaccccgaccgcctgtaaaagcgtgatgacaa							366
QY	378	GCAACAGCAACAGCTGTGTGAAGCTTTGAAGCGCATTTGGCTACACTGTAATCTTGCAAGAA							437
Db	367	gcccgcgcgcgccttggttgtaagagcttgtaaaacatcgcgcaaacctctgtaacattctaa							426
QY	438	GCTATTTGCTCAGGCTGCTGCAGTATCATCAGCTCT--GAAANAGTTGAGATATATC							494
Db	427	gaaatccgccaagctcgctctctatctccgcgaactctgcagcaacaagctgcgcgattatt							486
QY	495	TCGAGAAGCATATGGACGCTGTGGCAACGATGTGTGATATACATGGAAGAAATCTCGAGGT							554
Db	487	gccgaagcgaatggaanaaagtcgcgaagaagaagcgtgataacgcttgtaagagcgaacatct							546
QY	555	ATGGAACAGCAACTTGAAGTGTGTTCAAGCGATCAATTTGACCGTGATCTTACCTGCTCA							614
Db	547	ctggaanaaagagctggaagctggttgaaagttatgcgaattcgcgcgcggtctactgtctctct							606
QY	615	TACATGCTCACAGACAAATGAAATGTTGCGAGACCTTGAANACCATTTATCTTAATC							674
Db	607	tactcatcaacgaatgcygaanaaacaatacgcgcgcttggaacatccgttgttaltgcgt							666
QY	675	ACGGATAAAAAAGTGTCAACACATCCAAAGCATTTTGCCCACTCTTGAGAGAGTTCTTAAA							734
Db	667	ctcgcaaaaaaataatagcaatattctcgcgaccgcgtctgttltgtaaacaaatgtgccaat							726
QY	735	ACCACACCGCACTTACTCACTTAATTCGAGATGATGTGTGATGGTGAAGCACTTCCAACCCCT							794
Db	727	gccagccgcttcgctgttgaattatcgcttgtaagcgtatagaagcgcaagcctltgcygacttg							786
QY	795	GTCCTTGACAGAGATTCTGTGTACTTTCATATGTTGTTGCTGTCAAAAGCGCCAGAGATTTGGT							854
Db	787	gtcgtgaaacaacatccgcgcgcaattctgtaaaacgcttgcgtttaaagctccgcggtctgcgc							846
QY	855	GATTCGTCGTAAGGCTATGCTTGAAGACATTTGCTATCTTACACAGGTGTGACATGATTACA							914
Db	847	gaccgcgcgcaaaagcagtcgtcgaagacatcgcatacttcctacgcgcgcgcagctgatttcc							906
QY	915	GAGGATCTCTGAGCTTGAATTTAAAGATGCTTACATATGACAGAGCCCTTGAGCAGGCTCGTAAG							974
Db	907	gaagaagctcgcgctctcttggtaaaagcgaatttggacgacttgggtatagcgccaacgc							966
QY	975	ATTACAGTTGATAGATATGACAGATTAATGTTTGAAGGTTCTGAGAAATTCACAAACTAAT							1034
Db	967	atcgaataatcgtgaagaanaaacacacacatcatcgcgcgcttcgcgcgaacgcgcccaaatc							1026
QY	1035	GCTTAACCGTATTTGCACTGATTTAATCGCAATTTAGAAACACAACTTTCTGACTTTGACCGT							1094
Db	1027	gaagcgcgctgttgcgcaaatctcgcaacaacaaatcggaaacccgcgaacagcgtattacgacaa							1086

```
QY 1095 GAAAACTACAGAACCTTTGGCGAAATTAAGCTGGTGGTGTAGCTGTATCAAGTAGGA 1154
    ||||| ||||| || ||||| || ||||| || ||||| || ||||| ||
Db 1087 gaaaactgcaagacgcgttgcgaactgacgagcgcgctgggaatcaaatcgct 1146
QY 1155 GCTCCAAACAGACAGCTTTAAAGAAATGAACCTTGCAATGAGTGTCTAAATGCT 1214
    || || || || || || || || || || || || || || || || || ||
Db 1147 gccgcgacgcgaagtggaagaaagaaagacgcgtggaaagcgcgtgcacgct 1206
QY 1215 ACACGTCACACCCCTTGAGAGAGGTATGCTGCTGGTGGACAGCACTTATACGCTT 1274
    || || ||||| ||||| ||||| || ||||| || ||||| || ||||| ||
Db 1207 accgcgcgacgcgttgaaagagcgctgtgcagcgcgcgctgagccctgtgcgccc 1266
QY 1275 ATTGAAAATATACACAGCTTCTTGAGCTTG---AAGCGATGATGCTACTGACGTAACATT 1331
    || || || || || || || || || || || || || || || || || ||
Db 1267 cgtgcgcgtcttgaaaacccgcacacgcgcaatgcagcaagacgcgctacaaatc 1326
QY 1332 GTGCTTGCTGTAGAGAGCTGTACGTCAATGCTTTAAATGCTGGGTGGAAGGC 1391
    || || || || || || || || || || || || || || || || || ||
Db 1327 gctctgcgcgcgttgagctccgctgcgcgaatcgttcgaagcgcgcggaacc 1386
QY 1392 TCCGTAGTTATTTACAGATTGAAACAGCCCTGCAGAGAACGATTTAATGCTCAACA 1451
    || || || || || || || || || || || || || || || || || ||
Db 1387 agcgctgtgttgaaagaaagcttgaaagcaaatgctgttaccacgcctgcgacg 1446
QY 1452 GGTGAGTGGGTATATGATTAACAGATCATTTGACCTGTCAAGTAACAGATCA 1511
    || || || || || || || || || || || || || || || || || ||
Db 1447 ggcgaatacgcgcacatgcaatgagcgctactgcacccgcgcaaatgacccgttc 1506
QY 1512 GCGCTCAAAATGAGCTTGTAGTAGTCTTATTTTGACACAGAGAGATGTTGCT 1571
    ||||| || || || || || || || || || || || || || || || || ||
Db 1507 ggcgtcgaacagcgcgctcatalcgccgcgcgtatgctgcgacacagctcatalgct 1566
QY 1572 AATTAACCTGAACACAGCTACGCCAGCGCCAGCAATG 1607
    || || ||||| || || || || ||
Db 1567 gaaatccctgaagacaaacgcgctatgcctgatactg 1602

RESULT 10
X20248/c
ID X20248 standard; DNA; 910715 BP.
XX
AC X20248:
XX
DT 04-MAY-1999 (first entry)
DE Borrelia burgdorferi polynucleotide sequence #1.
XX
KM Borrelia burgdorferi; spirochete; bacterium; pathogen; Lyme disease;
KM epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;
KM infection; diagnosis; characterisation; detection; ds.
XX
OS Borrelia burgdorferi.
XX
PN W09858943-AL.
XX
PD 30-DEC-1998.
XX
PF 18-JUN-1998; 98WO-US12764.
XX
PR 03-SEP-1997; 97US-0057483.
PR 20-JUN-1997; 97US-0050359.
PR 22-JUL-1997; 97US-0053344.
PR 22-JUL-1997; 97US-0053377.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (MED-) MEDIMUNE INC.
XX
PI Clayton R, Dougherty BA, Fraser C, Lathigra R, Smith HO;
PI White OR;
XX
DR WPI: 1999-081217/07.
XX
```

```
PT New isolated Borrelia burgdorferi nucleic acids - used to develop
PT products for the detection, diagnosis, characterisation, prevention
PT and therapy of infections, particularly Lyme disease
XX
PS Claim 1: Page 157-671; 1128pp; English.
XX
CC X20248 to X20402 represent polynucleotide sequences isolated from
CC Borrelia burgdorferi (Bb). Products derived from Bb can be used for
CC the detection, diagnosis, characterisation, prevention and therapy of
CC Bb infections, e.g. Lyme disease. They can also be used for the
CC production of biosynthetic products, e.g. enzymes. Borrelia belongs
CC to a family of molle, spiral-shaped bacteria called Spirochetes.
CC Spirochetes are pathogenic in humans and Borrelia causes epidemic and
CC endemic relapsing fever, and Lyme borreliosis, more commonly known as
CC Lyme disease.
XX
SO Sequence 910715 BP; 327171 A; 129646 C; 130753 G; 323091 T; 54 other:

Query Match 33.2%; Score 551.6; DB 20; Length 910715;
Best Local Similarity 60.0%; Pred. No. 8.8e-134;
Matches 980; Conservative 0; Mismatches 639; Indels 15; Gaps 3;

QY 10 TTGATATGCAAAAAGAAATCAAAATTTTCACCAATGCGCGCTGCTGCATGCTGCGCGAG 69
    || ||||| ||||| || || || || || || || || || || || || || ||
Db 222236 TTTTATAGGCTAAAGACATATATTTTATAGAGATGCTAAGAAAAGCTTATGAGCGC 222177

QY 70 TTGATATGTTAGAGATACGCTCAAAAGTACGCTTGCTCTTAAAGCGGCAATGTTTCT 129
    || || || || || || || || || || || || || || || || || || ||
Db 222176 TTGAAAATTTATCCAAATGCTGTAAAGTAACTCTTGCGCCAAAAGSAGAAATGCTT 222117

QY 130 TTGAAAAGCTTTTGTGCTTCCTTAATTAATGATGAGGGGTACCATTTGCTAAAGAGA 189
    || || || || || || || || || || || || || || || || || || ||
Db 222116 TTGATATAAAAGTTGCGCTCCCAACGGTTACAAAGATGGGGTTAGCGTTCGTGAGA 222057

QY 190 TCGAATATAGAAGATCAATTTTGAAGAAACATGGGACAAATTTGGTCTGAAGTGGCTTCT 249
    || || || || || || || || || || || || || || || || || || ||
Db 222056 TTGAAGCTTGAATAATCCGTTTGAAGAACATGGGGACGCTTTTAAAGAAAGTTGCTT 221997

QY 250 AAACCAATGATATTTGCTGATGGATGGACGACATCTGCAACAGTTTGGACACAGCCATTG 309
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 221996 AAACCAATGATGTTGCTGATGGATGGACACACACTGCTACTGTTCTGCTATGCTATT 221937

QY 310 TTGATGAGGACTTAAATAATGTGACAGCAGGTGCTTAATCCAAATTTGATCCGTGAGGCA 369
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 221936 CAAGAGAGCGCTTAAAGAAATGCTCTTCAAGAAATCAATCCTATGGAATTAAGAGGAA 221877

QY 370 TTGAAACAGCAACAGCAACGCTGTGAAGCCTTGAAGCATTGCTCAACCTGATCTG 429
    || || || || || || || || || || || || || || || || || || ||
Db 221876 TAAGTACAGCTGTAAATTTGGCTGCTGAGAAATTTGCTGACTGTGCAAAAAGATTACAA 221817

QY 430 GCAAGAGAGCTATTTGCTCAGGTGCTGCAATATCATCAGCTCTGAAA---AAGTTGGAG 486
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 221816 CAAGAAGAGAGATTGCAACAGTGAATTTCTGTGTAATATGACAGTATATAGGTG 221757

QY 487 AGTATATCTAGAGACTATGAGAGCTGTGGGCAACATGCTGTGATTTACCATGGAAGAT 546
    || || || || || || || || || || || || || || || || || || ||
Db 221756 AAAAATTTGCTGAGGCAATGATTAAGTTGGAAAAGATGCTTTAATACAGTTGAAGAGT 221697

QY 547 CTCGAGGTATGAAACAGAACTGAAGTGTGTAAGGCAATGCAATTTGACCGGTGAAC 606
    || || || || || || || || || || || || || || || || || || ||
Db 221696 CAAAAACCTTTGATCTACGATTTCTTATGTTGAAGGTATGCAATTTGATGAGGATATC 221637

QY 607 TGTCTCAATACATGTCACAGACATGAAGAAATGTTGGACAGACTTGAAGAACCACTTTA 666
    || ||||| || || || || || || || || || || || || || || || ||
Db 221636 TTTCCTCCTTATTTTCTACCAATAAAGAAATATGAGTGTATTTTGGACGATGCTTTCA 221577

QY 667 TCTTAATTCAGGATTAAGAAAGTGTCAAAACATCCAGACATTTTGGCACTACTTGAGGAAG 726
    || || || || || || || || || || || || || || || || || || ||
Db 221576 TATTGATATATGAGAAAAGATAGTTCTATTAAGAGCTTTTACCAAGTCTTGAGAAAG 221517

QY 727 TTCTTAAACCAACGCTCATTTACTATTTATGCAATGATGATGGATGGTGAAGCACTTC 786
    || || || || || || || || || || || || || || || || || || ||
```

D	221516	TTTTAGGCAAAATAAACCTTTTAAATTAATTAATGCTGAGGATATTGAGGGGAGATGCTCTTG	221457
O	787	CAACCCCTTGCTCTTAACAAGATTCGTGTACTTTCAATGTGGTTCGTCTCAAAAGCCGAC	846
D	221456	CTGCTCTTGTTTAAACAGGGTTAGAGAGACTTTAAAGGATATGTGCATTAATTAATCTCTG	221397
O	847	GATTTCGATCGTCGTAAAGCCTATGCTTAAGACATTCGTATCTTCACAGGTGGTACAG	906
D	221396	GTTTTCGTGATAGCAGAAAACAATGCTTGAGGATATTGCATGCTTACCGGGGGTGTTT	221337
O	907	TGATTACAGAGATCTAGGACTTGAATTAATAAAGATGCTACAAAGACAGCCCTTGACAG	966
D	221336	TAAATCAGTAGAGACTTAAGGCTTACCTCTTGAGACAGTTGAATTTGACAACTTGGACAG	221277
O	967	CTGCTAAGATTACAGTTGATTAAGATATGACACAGTAATTTGTAAGGTTCCAGAAAGTTAC	1026
D	221276	CTAAACATATTAAAGTTGATTAAGAACATACCACTATTATTAA---TACGGCAATTAAG	221220
O	1027	AAGCTATTTCCTAACCCGTAATTCACATGATTAATTCGCAATTTAGAAACAACAACCTTCGACT	1086
D	221219	AGCAAAATTAAGAGAGCTTCAGAGCTTATTTAAAGCAAAATTTGAAGATTCACATCTGAAAT	221160
O	1087	TTGACCCGTGAATAACTACAAAGACGTTTGCGCAATTAAGCTGGTGGTATGCTGTAATCA	1146
D	221159	ATGATTAAGAAATAAATCTCAAGAGGCTCTTCGCAAACTTGTGGCGGAGTTGCTGTATTTA	221100
O	1147	AAGTAGAGACTCCAACAGAGACAGCTTTAAAGAAATGAATCACTTCGATTTAGAGATGCTC	1206
D	221099	ATGTTGGAGCTTACTAGAGGTAGAGCTTAAGAGAAAAGCATATGAGTTGAGAGAGCTC	221040
O	1207	TAAATGCTACACGTGCGAGCCGTTGAAGAAGTATCGTTGCTGGTGGTGGACAGCACCTTA	1266
D	221039	TTTCTGCAACTCGTCGCTGTTGAAGAAGGGGTGTGGCTGGCGGGTGATGTCAACTCTTA	220980
O	1267	T-----TACGTTATTGAAAATAAGCTAGACGCTTGTGACTTGAAGGGGATGATGCTA	1317
D	220979	TTGGAAGTTCGATATGATTTAGATATCAATATATACAGTAAATTTAAGCTATGAGAGAAAAGC	220920
O	1318	CTGAGACGTAACTATTTGCTCTGCTCTAGAGAAGCCGTGATGCTCAAAATTTGCTTTAAATG	1377
D	220919	AAGGTTTGGATGTATAAAGAAAGACTTGAAGAAGCCATGAGACGATTAATTTCAAATG	220860
O	1378	CTGGGTACGAAGGCTCCGCTACTTATTACAAAGTTGAAAAACACCCCTGCAGGAACAGGAT	1437
D	220859	CTGGTTTGAAGGATCTAATTTATTTATCATCAAAATTAACAGAAAAAAAAGGGCTTGGGT	220800
O	1438	TTATATGCGCAACAGAGGATGGGTTGATTAAGATTAACAGAGATCATGAGCCGCTCA	1497
D	220799	TTGATGCTCTCCAGCTTTTAAGTGGGTAATATGATATGATGAGCTGGAATTAATTTGATCTGCTA	220740
O	1498	AAGTAAACAGATTCAGCCCTTCAAAATCCAGCTTCTGTAGCTACTCTTATTTTACACAAG	1557
D	220739	AGGTTACAGAAGAGGGCTTCAAAATGCTGCTCAATTCGTGACATTTTATTACACAAG	220680
O	1558	AAGCAGTTGTTGCTAATTAACCTGAACCACTACGCCAGCGCCAGCAATGCCAGCAGGTA	1617
D	220679	AATGTGCAATACAGATATTAAAGAAAGAAAATAATCTCTGGTGGNGTGATTATCTCA	220620
O	1618	TGGATTCAGGAATG 1631	
D	220619	TGGAACCCAGGAATG 220606	

RESULT 11

ID	Q131.36	standard; DNA; 2465 BP

AC Q13136;

DT 22-OCT-1991 (first entry)

DE Hyp operon.

XX

KM Antibodies; heat shock; hypersensitive; allergen; HSP60; GROEL,
GROES, ss.
XX
OS Chlamydia psittaci GPIC.

FH	Key	Location/qualifiers
1E	000	014 000

```
/*tag= a
```

FT	673..2307
CDS	

```

EFT
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
```

```
/*tag= c
```

```
/*tag= d
```

```

/*tag= e

```

```

ET
/*tag= f

```

```
/*tag= g
```

```

F1  misc_feature  350..385

```

```
FT
/label = oligo
```

FT	misc_feature	2152..2187
----	--------------	------------

```
FT
/label = oligo
```

[illegible]

3
3
4
4
4
4

XX 0000 014500
00 0000 1000

XX
XX
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
8

XX
XX
ATTORNEY AT LAW
FITCHBURGH
MASSACHUSETTS

XX 1001 246603 233
DD EDT

DR P-PSDB; R133334.

3 XX

prods. for detection of and vaccines against C

PS Disclosure; Fig 5; 51pp; English.

CC The sequence was obt'd. from clone pGP57 isolat

CC HyPB hypersensitivity proteins, regulated by a

proteins or to design probes for the detection

in vaccines. The Hyp B protein elicits a cell

CC See also Q13137.

50 Sequence 2465 BP; 853 A; 467 C; 509 G; 636 T;

Cum gratia
Mortimer
T. J.
C.
E.
C.
J.
J.
J.

Best Local Similarity 60.38; Pred. No. 2.2e-13

10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044

Dh
670 ++++++
acaaaaaa+at+aaaata+aaaccccccccccc+

78 **თარგმანები**

Db 739 cttgcgaagaagctgtaaaagaaaccttaagtlccctaaagaagccgtacgtggtatccgaataa 798
 Oy 138 GCTTTTGCTTCCTCCCTTAATTAATGACGGGTAACCAATGGTAAAGAGATGAATTA 197
 Db 799 agcttggctcctcccaagttaccaaagatgagcgtaactgctgctaaagaatltgagctc 858
 Oy 198 GAAGATCATTTTAAACATGGAAGCAAAATTTGGTGTCTGTAAGTGGCTTTAAACCAAT 257
 Db 859 gaagacaacgcatgagacaatgaggaagctcaaatgtttaaagaagtcgtacgaacaatgca 918
 Oy 258 GAAATTTGCTGTGATGGAGCACTACTGCAACAGTTTGTGACACAGCACTGTTGATGAA 317
 Db 919 gataaagctgggtgagtaagaaactacacagctacgtctctgcaagaagcatcacaagtaa 978
 Oy 318 GGATTAATAAATGTGACAGCAGGTGCTTAATTCATTTGTTCCGTGAGCAATGGAACA 377
 Db 979 gggcttgaagaagcttaactgagagcgccaaatccctatgacctcaaaagaagcatgtataag 1038
 Oy 378 GCACAGCAACAGCTGTGAAAGCTTGAAGCCATTGCTCAGCACTGTATCTGGCAAGGAA 437
 Db 1039 gcaagtaaaagctgtgtcgtatgaaatcaaaaataatagtaaaacccgtacacacacaa 1098
 Oy 438 GCTATTTGCTAGTGGCTGCTGATCATCAGCTCTGA---AAAGTTGAGAGTATATC 494
 Db 1099 gaaatagctcaagctagacatcttctgcaataatgagatgctgaaatccgtatctatc 1158
 Oy 495 TCAGAGCTATGAGAGCGGTGGGCAACGATGTGTGATTAACATTCAGAAATCTCAGAT 554
 Db 1159 gcgcaagcatgtaaaagtlgtgcaaaaagcctctatctatgttgaagaagctaaagtc 1218
 Oy 555 ATGGAACAGCAACCTTGAATGGTTGAAGCATGCAATTTGACCGTGTACCTGTCTCAA 614
 Db 1219 ttccgaacatctccctgcagctgtgtcgaaggtatgaaatctcaacgcggtacaccctcagc 1278
 Oy 615 TACATGCTCAGACATGAAATGATGTGACAGCTTGAACCAATTTATCTTAATC 674
 Db 1279 tacttcttcaaaatccctgtaaaacaagaatgtgtttagaaagaagctcgtgttatc 1338
 Oy 675 ACGGATTAATAAAGCTGTCAAAATCAACATTTTGCACACTTGTGAGAAGTTCTTAA 734
 Db 1339 tatgataaaaaatctccgaaatcaaaagatttctacccagtttacaagaagtagaagaa 1398
 Oy 735 ACCAACCGTCCATTAATCAATTTATGAGATGATGTGATGATGATGATGATGATGATGAT 794
 Db 1399 tcaaggaagctccctacttacttctgtgaagatcagaaagaaagcctttagactttaa 1458
 Oy 795 GTCCTTAACAAGATTTGCTGATCTTCAATGTGTTGCTGTCAAAAGCGCAGGATTTGGT 854
 Db 1459 gtagttaaacaagactaagctgtgattcagaagtggtgcagtaaaagctccctgatttgc 1518
 Oy 855 GATTCGTGTAAGCTATGCTTGAAGACATTTGCTATCTTGAACAGGTGTACAGTATTACA 914
 Db 1519 gatatacaaaaagctatgttagaagacacgcgtactttaaactgtgtgtaacatccatcagc 1578
 Oy 915 GAGAGATCTGAGCTTAATTAATAAGATGCTACATGACAGCCCTTGGACAGGTGCTAAG 974
 Db 1579 gaagagctgtgagatgaaagcttgagaacacacacttagcatgttagaagaaagctaaaaa 1638
 Oy 975 ATTACAGTTGATTAAGATGATGACAGATTAATTTGTAAGTTGACAGGTTCAGAGGATTA 1034
 Db 1639 gtcacatgcttccaaagaataatacaaatgtttagaagctcctgagcaagaagaagatalt 1698
 Oy 1035 GCTAACCGTATTTGCATGATTAATTCGAATTAAGAAACAACAACCTTCTGACTTTGACCGT 1094
 Db 1699 gatctcgtcgtgaaagatatacaaaaaaacaatccgaagacagtaactctgatttagcctaaa 1758
 Oy 1095 GAAAAACTCAAGACAGCTTTGGCGAAATTAAGCTGGTGTGTGCTGTATTCAGAAAGTAGA 1154
 Db 1759 gaaaaactcccaagaagctttagactaaacttccgagagcgtagctgtaactcgtctaga 1818
 Oy 1155 GCTCAACAGACAGCTTTAAAGAAATGAACCTTCGATTTGAGGATGCTCTTAATTTGCT 1214

Db 1819 gctgtctacagaataatcgaaatgaaagaaaaagacagagtagatgctgacatgca 1878
 Oy 1215 ACAAGTCCACCCGTTGGAAGAAGGATGCTGCTGTGTGTGGAACAGCACTTAATTAACGTT 1274
 Db 1879 accttgctgcaagttgaagaaggtatctactacgtcggtgtgtacagcttagtctgctgc 1938
 Oy 1275 ATTGAATAAGTACAGCTCTTGAAGCTTG-----AGGGCATGATGCTACTGACGT 1325
 Db 1939 atccctactttagaagcttccatctctatcttactaacaatgaagatgagcaaatcgagca 1998
 Oy 1326 AACATTTGCTGTGCTGTGAGAAAGCCGTGACGTTGCTGCTTAATTTGATGCTGGGTAC 1385
 Db 1999 cgtatgtcttccaagaataltccgcctcaatlaagaacaatgcaagcaatgctgtgaa 2058
 Oy 1386 GAAGCCCTCCCTAGTATTTGACAAAGTTGAATAACACCCCTCAGAACAGGATTTAATGCT 1445
 Db 2059 gaagcgctatcaatctgtcaacaagtgcttccgcctccctagcgaaagctatgtatgc 2118
 Oy 1446 GCACAGGTGAGTGGTGTGATGATTAATAACAGAAATCAATTAACCTGTCAAGTAGA 1505
 Db 2119 ttacgcgattgtctacacgcagcatgtagagcaggaatcttcgactcaactaaagttaa 2178
 Oy 1506 CGATTCAGCGCTTCAAAATGCAAGCTGCTGATGCTAGTCTTATTTGACACAGAACAGT 1565
 Db 2179 cgttgcgtttagaaagcgagcttctgtagctg99cttctatlaacaacagaagctta 2238
 Oy 1566 GTTGTAAATTAACCTGAA 1583
 Db 2239 atgcccagatattcctgaa 2256
 RESULT 12
 A13000
 ID A13000 standard; DNA; 1635 BP.
 AC A13000;
 XX
 DT 15-AUG-2000 (first entry)
 XX
 DE Neisseria meningitidis gene sequence SEQ ID NO:2964.
 XX
 KW Neisseria meningitidis; Neisseria gonorrhoea; antigen; vaccine;
 KM antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
 KW antibacterial; gene therapy; ds.
 XX
 OS Neisseria meningitidis.
 XX
 PN M09957280-A2.
 XX
 PD 11-NOV-1999.
 XX
 PF 30-APR-1999; 99WO-US09346.
 XX
 PR 01-MAY-1998; 98US-0083758.
 PR 31-JUL-1998; 98US-0094869.
 PR 02-SEP-1998; 98US-0098994.
 PR 02-SEP-1998; 98US-0099062.
 PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103794.
 PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 99US-0121528.
 XX
 PA (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 PI Petersen J, Pizze M, Rappelli R, Ratti G, Scatato E, Scarselli M;
 PI Tettelin H, Venter JC;
 PI
 DR WPI: 2000-062150/05.
 XX
 PT Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics

XX Neisseria meningitidis.
 OS
 XX MO9957280-A2.
 PN
 XX
 PD 11-NOV-1999.
 PF 30-APR-1999; 99MO-US09346.
 XX
 PR 01-MAY-1998; 98US-0083758.
 PR 31-JUL-1998; 98US-0094869.
 PR 02-SEP-1998; 98US-0098994.
 PR 02-SEP-1998; 98US-0098962.
 PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103794.
 PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 99US-0121528.
 XX
 PA (CHIR) CHIRON CORP.
 PA (GENO) INST GENOMIC RES.
 XX
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 PI Petersen J, Piza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX
 XX WPI: 2000-062150/05.
 DR P-PSDB: Y75746.
 PT Novel Neisserial polypeptides predicted to be useful antigens for
 XX vaccines and diagnostics -
 XX
 PS Claim 7; Page 1385-1386; 1453pp; English.
 XX
 CC 253015 to 254536, 254577 to 254615, and Y74253 to Y75941 represent
 CC novel Neisseria meningitidis and N. gonorrhoeae polynucleotides and
 CC polypeptides. 254537 to 254576 and 254616 to 255473 represent PCR
 CC primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
 CC presence of Neisseria bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.
 CC
 CC Sequence 1635 BP; 449 A; 435 C; 451 G; 300 T; 0 other;
 SQ

Query Match 32.3%; Score 536.8; DB 21; Length 1635;
 Best Local Similarity 59.4%; Pred. No. 5, 2e-131;
 Matches 948; Conservative 0; Mismatches 642; Indels 6; Gaps 2;

QY 18 GCAAAAGAAATCAAAATTTTCAGCAGATGCGCGTCGTCATGGTGGCGGAGTTGATAG 77
 DB 7 gcaaaagaaacagtcagtcgcaatgaagtcgcaaaaatgtaaaacgctggaacatt 66
 QY 78 TTAGCAGATACCGTCAAGTAAAGCTTGGTCTTAAGGCGGCAATGTTGTTTGAANA 137
 DB 67 ctgagcaaacgctgcgcgttaacctgggcccacaaagtcgcgaacgtaactgctgcgcg 126
 QY 138 GCTTTTGTTCTCCCTTAATTAATGACGGGGTAACCATTTGCTTAAGAGATGAATTA 197
 DB 127 gacttcgagcgcccgacacatcaacaaagcgcgttaaccgtccaaagaatcgaaactg 186
 QY 198 GAAGATCATTTTGAACAACATGGAGCAAAATGCTGCTGAAGTGGCTTTAAACCAAT 257
 DB 187 aaagacaagcttgaataatctggcgcaacatgtaaaagagctgctlccaaacaaac 246
 QY 258 GATATTTGCTGATGATGGAGCACTACTGACACAGTTTGAACAAACGATTTGTTGAA 317
 DB 247 gacgtggcgagcgagcgatcgactaccgcaacgctactggtgcaatccactgctgcgaa 306

QY 318 GGAATAAAAATGTGACAGCAGGTGCTTAATCCAATTTGGTATCCGTCGAGCATTTGAACA 377
 DB 307 ggtatgaatatatgttaccagagtgatgaatccgacgactgaacgcgttatcogataaa 366
 QY 378 GCAACAGCAACAGCTGTTGAAGCCTTGAAGCCATTGCTCAACCTGATTTGGCAAGAA 437
 DB 367 gccgtcgccgttctgtgttgaacacttgaacacatcgccaaactcttgcgaacttcaa 426
 QY 438 GCTATTGCTCAGTGGCTGAGTATCATCACGCTC--TGAAAAGTGGAGAGTATATC 494
 DB 427 gaatgcccaagtcgctctatcttcgccaactcgcgaacaaatcggcgcgatattac 486
 QY 495 TCAGAACCTATGAGCGCTGTGGCAACGATGTTGATTAATCAATCGAATCTCGAGGT 554
 DB 487 gccgaagcgaatgaaaaaagtcggaagaagcgatgaactacgttgaagcgcgaagctc 546
 QY 555 ATGGAACAGCACTTGAACGTGTTGAAGCATCAATTTACCGTGTACCTGCTCAA 614
 DB 547 ttgaaaacgagctggagcgtgattgaagatcagcttcgacgcggtacactgtctcct 606
 QY 615 TACATGCTCACAGACATGAAAAATGTTGCGAGACCTTGAAAAACCATTTATCTTAATC 674
 DB 607 tactcatcaacgactcggaaaaaaacaatccgcgttggacaacccgttgcattcgttg 666
 QY 675 ACGGATAAAAAAGTGTCAACATCAAGACATTTTGCCACTACTTGAAGAGTTCTTAA 734
 DB 667 ttgcgaacaaaaatcagcaacatccgcgactctgcgttcttggaaacagtgycgaaaa 726
 QY 735 ACCAACCCTTCATTACTATTATTCAGATGATGTGATGTCGACGACTTCCACCCCTT 794
 DB 727 gccagccgttcgctgttgcattatcgtgaagcgtgaagagcgaaacgttcgagacttg 786
 QY 795 GTTCTAACAGAAATGCTGCTGCTTCAATGTTGCTGCTGCTGCTGCTGCTGCTGCTG 854
 DB 787 gtctgtgaacacatccgagtgatcctcgtgaacacgcttgcgtcaaacccctgtctgcgc 846
 QY 855 GATGCTGTAAGACTATGCTTGAAGCATTTGCTTATCTTGAAGTGTACAGTGAATTA 914
 DB 847 gaccgcgcaaacgagtggttgcagacatccatccgtcgcgcggttcgtggtatcttc 906
 QY 915 GAGCATCTAGAGCTTAATTAATTAAGATGCTACATACACCCCTTGGACAGCTGCTAAG 974
 DB 907 gaagagctgctgtcttcttcttgaagaaagcacttggagacttgggttcaagcaaacgc 966
 QY 975 ATTACAGTTGATTAAGATAGACAGATTAATGTTGAAGTTCAGAGTTCAGAACTATT 1034
 DB 967 atcgaatatcgttaagaataaacaccacatcatcgaacgcttggcgaaacgacccaatc 1026
 QY 1035 GCTAACCGTATTGCTGCTGATTAATGCAATTAAGAAACAACATTTGACTTTGACCGT 1094
 DB 1027 gaagcgcggttgcgaaatccgcgaacatccgaacacgcaacgacgattacgcgaac 1086
 QY 1095 GAAAACTACAGACAGCTTTGGCGAAATTAAGTGTGCTGCTGCTGCTTATCAAGTAGA 1154
 DB 1087 gaaaactctgaagagcggttgcataatctggcagcggttgcagcttcaaaagtcgct 1146
 QY 1155 GCTTCAACAGAGACAGCTTTAAAGAAATGAACCTTCGATTTGAGAGTCTCTAAATGCT 1214
 DB 1147 gccgcgaacggaatcgaatgaagaagaaagaaagacgcgcgttgaagacgcgcgcgcgc 1206
 QY 1215 ACACTGACGCGCTTGAAGAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1274
 DB 1207 accgcgcgcgcgttgaagaagagcggttgcagcggttgcagcggttgcagcggttgcgcgc 1266
 QY 1275 ATTGAAAAAGTACAGCTTTGAGCTTG--AGGCGATGATGCTACTGAGAGTAACTAT 1331
 DB 1267 cgttgcgttcttgaagaacacccacacccgcgaatcgcgaacgaagacgcgcgttacaac 1326
 QY 1332 GTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1391
 DB 1327 gcttgcgcgcgcgttgcagcttccgcgtcgcgcgaacatcgttgcgaacgcgcgcgcgcgc 1386

Oy		1392 TCCGTAAGTATTATGCACAAAGTTGGAAAAAACAGCCTCGCACGGACAGCATTTAATGTCAACA	1457
Oy			
Dd	1387 agcgttgcttgcgaacaagatttcgaaaggcaaacgaactagaacgctcgccacc	1448	
Oy	1452 GGTAGTAGGGTGTTATGATTTAAAACAGGAATCATTCGACTCGTCAAAGTACAGATCA	1511	
Dd	1447 gggcgataacgccatgatcatcgaaatggcgtacltcgacctccgccaagttaaaccgcttc	1506	
Oy	1512 GCAGTTCAAAATGCAGCCTTCCTGAGCTAGTCTATTATTGACAACAAGAACAATTGTTGCT	1571	
Dd	1507 gcgctgaaaacaacggcgatcatacgcgcgcttgatgctcgaccaacatcgatgcatacgct	1566	
Oy	1572 AATAACCCTGACCACGCTACGCCAGCCAGCCAAGCATG	1607	
Dd	1567 gaatccccggaagaacaaccgctgtgcctgatatg	1602	
 RESULT_14 O47926 Q47926 standard; DNA; 1838 BP.			
XX AC	O47926;		
XX DT	16-MAR-1994 (first entry)		
DE DE	hsr60 DNA.		
XX XX	H. pylori; heat shock protein; hsr60; conjugate compound; cellular; immune system; antibody; bacteria; H. influenzae; Streptococcus; Salmonella; vaccine; Meningococci; ss. 		
OS OS	Helicobacter pylori.		
FH FT	CDS Location/Qualifiers		
FT CDS	58..1698 /*tag= a		
FT RBS	/product= hsr60 47..51		
XX FT	/*tag= b		
PN PN	M09317712-A.		
XX PD	16-SEP-1993.		
PF PF	08-MAR-1993; 93MO-EPO0516.		
PR PR	06-MAR-1992; 92IT-OFI0058.		
PA PA	(BIOC-) BIOCIINE SCLAVO SPA.		
PI PI	Costantino P., Norelli F., Rappunni R:		
DR DR	WPI: 1993-303147/38.		
XX Xx	p-PsDB; K41661.		
PT PT	New conjugates of heat shock protein and oligo- or poly:saccharide		
PS PS	- used in vaccines or to prevent or treat bacterial infection		
Xx xx	Disclosure; Fig 3; 69pp; English.		
CC CC	This sequence encodes the H. pylori heat shock protein, hsr60.		
CC CC	The protein encoded by this sequence may be used as part of a		
CC CC	conjugate compound, along with at least one oligo- or polysaccharide.		
CC CC	The hsps are highly conserved across bacteria and they stimulate the		
CC CC	cellular immune system. When they are conjugated to a polysaccharide		
CC CC	they provide an immunostimulatory effect and produce anti-poly-		
CC CC	saccharide antibodies in the absence of adjuvants and of pre-		
CC CC	sensitisation. The conjugates may be used as vaccines for		
CC CC	prophylactic or therapeutic purposes. They may esp. be used against		
CC CC	bacteria such as H. influenzae, Streptococcus, Salmonella and		
XX Xx	Meningococci.		
Xx xX	Sequence 1838 BP; 601 A; 345 C; 468 G; 424 T; 0 other;		

Query Match	31.88;	Score 528.6;	DB 14;	Length 1838;
Best Local Similarity	58.98%;	Pred. No. 7.7e-129;		
Matches 966;	Conservative	0;	Mismatches 664;	Indels 9;
			Gaps	
QY	13	ATATGGCAAAAGAAATCAATATTTTCAGCAAGATCGGTGCTCCGATCGGTGGCCGGAGTTG	72	
Db	56	aaatgycgaagaatacaatttttcgataibgycggaacacttlatltgaagcgiga	115	
QY	73	ATATGTAGCAGATACCGTCAAGTAAGCGTTGGTCTTAAGGGCGCAATGTTGTTCTTG	132	
Db	116	gcaactccatgacgcgtctcaagtaaccatgycgcaagagcagaaatgatattgacc	175	
QY	133	AAAAAGCTTTGGTTTCCTCTTAATTACTATACGCGGTAACCACTTGCTAANAGATCG	192	
Db	176	aaaaagctatgycgcctccaaagcaaccacaagaacgycgagcgcggtctaaagatitg	235	
QY	193	AATTAGAGATCATTTTGAAGAACATGGGAGCAAAATTTGGTGTCTGAAGTGGCTTTAA	252	
Db	236	aattaaigtgccagtagcttaacatgycgtccactcgtttaagaagtagagcaaa	295	
QY	253	CCATGATATTGCTGGTGGATGGGACGACTACTCAACAGTTTGGACACAGCACTTGTTC	312	
Db	296	ccgctgtagcgcgcgagcgaibgcagacccacgacgcgtgtagcttaagactta	355	
QY	313	ATGAGAGCATTAATAATGTACAGACAGGTCTTAATCCATTGGTATCCGTGACGACATTG	372	
Db	356	aagaagatltgaagaaataatacaagcgcggygcctaaccttbaagtgaagaagcagc	415	
QY	373	AAACAGCAACAGCAACAGCGTTTGAAGCCTTGAAGCCATTGCTCAACCTGATATCGCA	432	
Db	416	ataaagcgtcgaagcgcatacttaatbagctttaaagaagcagcaaaaagtagcgcta	475	
QY	433	AGGAAGCATATTGCTCAGGTGGCTGCGAGTATCATCACGCTC---TGAAAAGTTGGAGAT	489	
Db	476	aagaagaatlaocgaagibgcagacattctgcgaactccgatacaaatatcgygaac	535	
QY	490	ATTATCTAGAAGCTATGTAGACGCTGTGGGCAACATGTGTGATTCATCATCAAGAAATCTC	549	
Db	536	tcactcgcagacgtatgtaaaaagibggtgaagaacgscgtgataccglttagagagcta	595	
QY	550	GAGTGTATGGAACAGCACTTGAAGTGGTGAAGCATGCAATTTGACCGTGGTAACTGCT	609	
Db	596	aaggcatatgaagatgaatibgatgctcgaagaagcatatgatalagaagctaccct	655	
QY	610	CTCATATCATGTGTACAGACAAATGAAAATGGTTGCAGCCGTGAAAACCCATTATCT	669	
Db	656	ccccttatltgtgaagaaocgctgagaaatgaacgcgtcaatitgatatgtttacacc	715	
QY	670	TAAATACGGATTAATAAAGTGTCAACATCCAGATCAATTTGGCACACTTGTAGAAAGTTC	729	
Db	716	ttttaacgataaaaaaatctctagcatgaaagacattctccgcctactagaaaaaacca	775	
QY	730	TTAAACCAACCCGTCATTTCTCATTAATTGCAATGATGATGGTGAAGCACTTCCAA	789	
Db	776	tgaagaagggcaaacgcgttlltaactcgcgcgaagacatigaagycgaagatlltaagca	835	
QY	790	CCCTGTGCTTGGAAACAAATTCGTGTACTTTCATGTGGTTCGTGCAAGCGCCAGAT	849	
Db	836	ctctagtggtgataaataatlaagagcggtigtgataatactgcagcgyttaaagctccagct	895	
QY	850	TTGGTGATCGTCGTAAGAAGCTATGCTTGAGACATTTGCTATCTTGAACAGTGGTACATGA	909	
Db	896	ctgaggaacagaagaagaatgctccaaagacatcgtctatlltaacccgycgltcaagicta	955	
QY	910	TTTACAGAGAGCTGTGAGACTTGAATTTAAAGATCTTCAATGACAGACGCCCTTGACAGGCTG	969	
Db	956	tttagcgaagaatibgycgtctagacttagaaacgcgcgaagtvgaglttttlaagcaagctg	1015	
QY	970	CTAAGATTACAGTTGATTAAGATAGCACAGTAATTTGTAAGAGTTTCAGNAAGTTTACAG	1029	
Db	1016	gaagatctgtatgtatgaagaagacacacccacgactctaataibgcgaagacatagccta	1075	

[illegible]

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2001, 20:54:05 ; Search time 71.35 Seconds
(without alignments)
3751.749 Million cell updates/sec

Title: US-09-001-737-7
Perfect score: 1661
Sequence: 1 GAATCGCTTCATATGCA.....TGGCGGATAGCCGATTC 1661

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 280836 seqs, 80580151 residues

Total number of hits satisfying chosen parameters: 561672

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents.NA.*
1: /cgn2_6/prodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/prodata/2/ina/6_COMB.seq.*
4: /cgn2_6/prodata/2/ina/PTUS_COMB.seq.*
5: /cgn2_6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	528.6	31.8	1838	3	US-08-470-260-7
2	528.6	31.8	1838	3	US-08-471-491-7
3	528.6	31.8	1838	3	US-08-466-662-7
4	496.6	29.9	2284	2	US-08-467-822-28
5	446.8	26.9	1626	2	US-08-997-080-159
6	446.8	26.9	1626	2	US-08-997-362-159
7	446.8	26.9	1626	3	US-09-095-855-159
8	444.8	26.8	1569	2	US-08-997-080-113
9	444.8	26.8	1569	2	US-08-997-362-113
10	444.8	26.8	1569	3	US-09-095-855-113
11	416.6	25.1	1620	2	US-08-461-775-10
12	416.6	25.1	1620	3	US-09-031-606-10
13	411	24.7	2668	2	US-08-461-775-11
14	411	24.7	2668	3	US-09-031-606-11
15	362	21.8	1320	2	US-08-461-775-8
16	362	21.8	1320	3	US-09-031-606-8
17	353.2	21.3	2167	2	US-08-461-775-9
18	353.2	21.3	2167	3	US-09-031-606-9
19	273.8	16.5	985	2	US-08-997-080-161
20	273.8	16.5	985	2	US-08-997-362-161
21	273.8	16.5	985	3	US-09-095-855-161
22	271.8	16.4	927	2	US-08-997-080-116
23	271.8	16.4	927	2	US-08-997-362-116
24	271.8	16.4	927	3	US-09-095-855-116
25	178.4	10.7	568	3	US-08-714-918-51
26	176.6	10.6	647	2	US-08-997-080-115
27	176.6	10.6	647	2	US-08-997-362-115
28	176.6	10.6	647	3	US-09-095-855-115

```

29 101.2 6.1 343 2 US-08-473-020A-7
30 100.4 6.0 337 2 US-08-997-080-77
31 100.4 6.0 337 2 US-08-997-362-77
32 100.4 6.0 337 3 US-08-873-970-77
33 100.4 6.0 337 3 US-09-095-855-77
34 97.4 5.9 350 1 US-08-105-168B-22
35 97.4 5.9 350 2 US-08-698-948-22
36 94.8 5.7 342 2 US-08-473-020A-31
37 94.8 5.7 343 2 US-08-473-020A-4
38 91.8 5.5 343 2 US-08-473-020A-2
39 91.6 5.5 342 2 US-08-473-020A-28
40 91.6 5.5 342 2 US-08-473-020A-30
41 91.6 5.5 343 2 US-08-473-020A-3
42 90.2 5.4 342 2 US-08-473-020A-29
43 90 5.4 360 2 US-08-997-080-79
44 90 5.4 360 2 US-08-997-362-79
45 90 5.4 360 3 US-08-873-970-79

```

ALIGNMENTS

```

RESULT 1
US-08-470-260-7
; Sequence 7, Application US/08470260
; Patent No. 6077706
; GENERAL INFORMATION:
; APPLICANT: Covacci, Antonello
; APPLICANT: Bugnoli, Massimo
; APPLICANT: Telford, John
; APPLICANT: Macchia, Giovanni
; APPLICANT: Rappuoli, Rino
; TITLE OF INVENTION: Helicobacter Pylori Proteins Useful
; TITLE OF INVENTION: for Vaccines and Diagnostics
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,260
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,848
; FILING DATE: 21-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0316.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1838 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-470-260-7

```

Query Match 31.8%; Score 528.6; DB 3; Length 1838;
Best Local Similarity 58.9%; Pred. No. 1.4e-146;

Matches 966; Conservative 0; Mismatches 664; Indels 9; Gaps 3;

QY 13 ATATGCCAAAAGAAATCAAAATTTTCAGAGATGCGCGTCTGCCATGTCGCGGAGTTG 72
 Db 56 AAATGCCAAAAGAAATCAAAATTTTCAGATAGTCGAGAAACCTTTTATTTGAAGCGTGA 115
 QY 73 ATATGTTACAGATACCCGTCAAAGTAAACCTTGGTCTTAAGGGCGCAATGTTGTTCTG 132
 Db 116 GGCACATCCATGACGCTGTCAAAGTAAACCTTGGTCTTAAGGGCGCAATGTTGTTCTG 175
 QY 133 AAAAGCTTTTGGTCTCCCTTAATTAATTAAGGGGTAAACCTTGGTCTTAAGGGGTAA 192
 Db 176 AAAAGCTTTTGGTCTCCCTTAATTAATTAAGGGGTAAACCTTGGTCTTAAGGGGTAA 235
 QY 193 AATTAGAAGTCAATTTTGAAGACATGGGAGCAAAATTTGGTCTGAAGTGGCTTCTTAAA 252
 Db 236 AATTAGAAGTCAATTTTGAAGACATGGGAGCAAAATTTGGTCTGAAGTGGCTTCTTAAA 295
 QY 253 CCAATGATATGCTGGTGAAGGAGCACTTACGACAGTATGACAAAGCCATTTGACAAAG 312
 Db 296 CCGCTGATGCTCGCGGATGTCGAGGACGACACAGGACCGTGTAGCTTATAGCATTTT 355
 QY 313 ATGAAGGACTAAAATGTGACAGAGTGTCTAATCCATTTGGTCTCCGTCGAGGATG 372
 Db 356 AAGAAGTTTGAAGATATACGCGTGGGCTTAACCTTATGAAGTGAACAGGAGCGATG 415
 QY 373 AAACAGCAACAGCAACAGCTGTTGAAGCCTTGAAGACCAATTTGCTCAACCTGTATCTG 432
 Db 416 AATAAGCTCTGAAGGATCAATTAATGAGCTTAAAGAGCGAGCAAAAGTAGCGGTA 475
 QY 433 AGGAAGCTATTCCTCAGTGGTGGTGAAGTGTGAGTATCATACGCTC---TGAAGAGT 489
 Db 476 AAGAAGAAATCAACCAAGTGGGAGCAATTTCTGCAAACTCCGATCAATATCGGGAAC 535
 QY 490 ATATCTAGAAGCTATGAGGCTGTGGGCAAGTGTGATGATGATGATGATGATGATGAT 549
 Db 536 TCATCGCTGACCTATGGAAGAGTGGGTAAAGAGCGGCTGATCAGCTGTTGAGGAGCTA 595
 QY 550 GAGGTATGGAACAGAACTTGAAGTGGTGAAGGATGCAATTTGACCGTGGTATACCTG 609
 Db 596 AGGCAATTGAAGTGAATTTGATGTCGTAGAGGATGCAATTTGATGAGGCTACCTCT 655
 QY 610 CTCATATGATGTCACAGCAATGAAAATGTTGTCAGACCTTGAAGACCCATTTATCT 669
 Db 656 CCGCTTATTTGTAAGCAAGCTGAGAAATGACCGCTCAATTTGATATGTTACATCC 715
 QY 670 TAATCAGGATAAAAGTGTCAACATCCAGACATTTTGCCTACTTGTAGGAGTTC 729
 Db 716 TTTTACGGATAAAAGTGTCTAGCATGAAAGACATTTCCCGCTACTAGAAAACCA 775
 QY 730 TTAACCAACCGTCCATTTACTTATTTGCAAGATGATGTTGATGTTGATGTTGATGTT 789
 Db 776 TGAAGAGGCAACCGCTTTTAAATCATCGTGAAGACATTTGAGGCGAAGCTTTAACA 835
 QY 790 CCGTGTCTTGAACAGATTCGTGTACTTTCAATGTTGTTGCTGTCAAGCGCCAGAT 849
 Db 836 CTCATGTTGTAATTAAGAGCGTGTGATATGCGAGCGTTAAAGCTCCAGCT 895
 QY 850 TTGTTGATGCTGTAAGCTATGTTGAAGACATTTGATCTTTGACAGTGTGTACAGTCA 909
 Db 896 TTGGGACAGAAAGAAATGCTCAAGACATCGCTATTTTAAACCGCGCTCAAGTCA 955
 QY 910 TTACAGGATCTAGGACTTGAATTTAAAGATGCTTACATGACAGCCCTTGGACAGGCTG 969
 Db 956 TTAGCGAAGATTTGGCTTGAAGTCTAGAAAAGCGTGAAGTGGAGTTTTTAGCGAAGCTG 1015
 QY 970 CTAAGATTTACATTTGATTAAGATACCAAGTAAATTTGTAAGGTTTCAGAAAGTTCAG 1029
 Db 1016 GAAGGATTTGATTTGACAAAGCAACACCGATCGTAGATGGCAAGGCCATAGCGATG 1075
 QY 1030 CTATTGCTAACCGTATTCAGCTGATTAATCGCAATTTAGAAACACAACTTCTGACTTG 1089
 Db 1076 ATGTTAAAGACAGATCGCGGAGATCAAAACCCAAATTTGCAAGTACGACAAAGCGATT 1135

QY 1090 ACCGTGAAAAAAGTCAAGAACGTTTGGCGAAATTAGCTGGTGGTAGCTGTATCAAAAG 1149
 Db 1136 ACAAGAAAAATTTGCAAGAAAGATTGGCTAAACTCTCTGGCGTGGCTGTGATTAAAG 1195
 QY 1150 TAGGAGCTCCACAGACAGAGCTTTTAAAGAAATGAAATTCGCATTGAGGATGCTCTAA 1209
 Db 1196 TGGGCGTCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1255
 QY 1210 ATGCTACAGCTGACGCGTTGAAGAGGATGATGCTGCTGCTGCTGCTGCTGCTGCTG 1269
 Db 1256 GCGGCACTAAAGCGGCTTGAAGAGGATGTTGATGTTGGTGGCGTGGCTCTCATTC 1315
 QY 1270 CGGTTATTTGAAAAAGTAGCAGCTTCTGAGCTTGAAGGCGATGATGCTGACGCTAA 1329
 Db 1316 GCGGCGCTCAAAAGT---GCATTTTCAATTTGCACGATGATGAAAGTGGGCTATGAA 1372
 QY 1330 TTGTCGCTGCTCTAGAGAGCGCTGTACGTCAAAATGCTTTTAAATGCTGGGTAGCAAG 1389
 Db 1373 TCATCATGCGCGCAATTTAAAGCCCAATTTAGCTCAAAATGCTTCAACGCTGTTATGATG 1432
 QY 1390 GCTCGCTAGTTTACAGAGTTTGAAGAAACAGCCCTGCAGGAACAGGATTTAATGCTGCA 1449
 Db 1433 GCGGTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1492
 QY 1450 CAGGTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1509
 Db 1493 ATGCAAGTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1552
 QY 1510 CAGCGCTCAAAATGCAAGTCTCTGAGTGTGCTTATTTTGAACAGAGAGGATGTTG 1569
 Db 1553 TCGCTCTCAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1610
 QY 1570 CTAATAACCTCAACAGCTACGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1629
 Db 1611 -CCATGAAATCAAGAAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1669
 QY 1630 TGATGGTGGGATGGCGG 1648
 Db 1670 GTATGGGAGGATGGCGG 1688

RESULT 2

US-08-471-491-7
 ; Sequence 7, Application US/08471491B
 ; Patent No. 6090611
 ; GENERAL INFORMATION:
 ; APPLICANT: Covacci, Antonello
 ; APPLICANT: Bugnoli, Massimo
 ; APPLICANT: Telford, John
 ; APPLICANT: Macchia, Giovanni
 ; APPLICANT: Rappuoli, Rino
 ; TITLE OF INVENTION: Helicobacter Pylori Proteins Useful For Vaccines And
 ; TITLE OF INVENTION: Helicobacter Pylori
 ; FILE REFERENCE: CHIR0044
 ; CURRENT APPLICATION NUMBER: US/08/471,491B
 ; CURRENT FILING DATE: 1995-06-06
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 7
 ; LENGTH: 1838
 ; TYPE: DNA
 ; ORGANISM: Helicobacter pylori
 US-08-471-491-7

Query Match 31.8%; Score 528.6; DB 3; Length 1838;
 Best Local Similarity 58.9%; Pred. No. 1.4e-146;
 Matches 966; Conservative 0; Mismatches 664; Indels 9; Gaps 3;
 QY 13 ATATGCCAAAAGAAATCAAAATTTTCAGAGATGCGCGTCTGCCATGTCGCGGAGTTG 72
 Db 56 aaatggcaaaagaaatcaaaattttcagatagtcgagaaacccctttttatttgagcggtga 115

QY 73 ATATGTTAGCAGATACCGTCAAGTAACGCTTGGTCTTAAGGGCGCAATGTTGTTCTTG 132
 Db 116 ggaactccatgacgctgtcaagtaacacatgggccaagagcgaggaatgtattgatcc 175
 QY 133 AAAAGCTTTTGTCTCCCTTAATTAATCAATGACGGGTAAACCAATGCTTAAAGAGATCG 192
 Db 176 aaaaagctatggcgtcccaagcatcaccaaaagacggtcgagcggtgcaaaagagatg 235
 QY 193 AATTAGAAGCATATTTTGAAGACATGGGACAAATTTGGTGTCTGAAGTGGCTTCTAAAA 252
 Db 236 aattaagtgcagtagctaaactaagcgctcaactcgttaagaagtagcgagcaaaa 295
 QY 253 CCAATGATATTTGCTGTGATGGGACGACTACTGCACAGTTTGTGACACAGCCATGTTTC 312
 Db 296 ccgctgatgctgcggtgagtcgacgacacagcgccgtgctagcttatagcatcttta 355
 QY 313 ATGAAGSACTAAAAATGTGACAGCAGGTGCTAATCCAATGTTGATCTCCGTCGAGGCAATG 372
 Db 356 aagaaggttgaggaataatacagcgctggggctaacctctattgaagtgaacgaggtcatg 415
 QY 373 AAACAGCAACAGCAACAGCTGTTTGAAGCCTTGAAGCCATTGCTCAACCTGTATCTGGCA 432
 Db 416 ataaagctgctgaagcagatcatttaagtgcgttaaaaaagcgagcaaaaagttagcggtg 475
 QY 433 AGGAAGCTATTTGCTAGGTCGCTGCTGAGTATCATACGCTC---TGAAAAAGTTGGAGAGT 489
 Db 476 aagaagaatacaccagtggtgcgaacctcttctgcaactccgcatcacaatatcggtgaac 535
 QY 490 ATATCTCAGAAAGTATGGAGCGTGTGGGCAACGATGGTGTGATTCACATCGAAGATCTC 549
 Db 536 tcctgctgacgctatggaaaaagtggtgaaagcgcgtgataccgcttgaggaagcta 595
 QY 550 GAGGTATGAAACAGAACTTGAAGTGGTGAAGCCATGCAATTTGACCGTGGTTACTCTGT 609
 Db 596 agggcaattgaagatgaattgagtgctgtagaaggaatgaattgtagaggtactctct 655
 QY 610 CTCATATACATGTGCAGACAAATGAAAAATGTTGTCAGACCTTTGAAACCCCATTTATCT 669
 Db 656 ccccttatttgaagcaagcgtgagaaatgaccgctcaattggataatgcttatatcc 715
 QY 670 TAATCAGCATAAAAAGTCTCAACATCCAGACATTTTGCACACTACTTGAGGAAGTTC 729
 Db 716 tttaacgataaaaaactctagcatgaagacattctcccgctactagaaaaacca 775
 QY 730 TTAACACCAACCCCTCCATTACTCATTTATGCAGATGATGGATGGTGAAGCACTTCCAA 789
 Db 776 tgaaagagggcaaacgcgttttaatacatcgctgaagacattgagggcgaaagctttaacga 835
 QY 790 CCCTTCTTGAACAGATTCGTGGTACTTTCAATGTTGTTGCTGTCATCAAGCCCGAGGAT 849
 Db 836 ctctagtggtgaataaataaagggcggtgtgaatatcgagcggttaagctccaggtc 895
 QY 850 TTGGTATGCTGCTTAAGCTATGCTTGAAGACATTCGCTATCTTGACAGGTGGTACAGTGA 909
 Db 896 ttgggacagaagaagaataatgctcaagacatcgctattttaaccggtcggtcaagtc 955
 QY 910 TTAACAGAGATCTAGGACTTGAATTTAAAGATGCTACAATGACAGCCCTTGACAGGCTG 969
 Db 956 ttgagcaagaattgggcttgagctgtagaagaacgcgtgaagtggagtttttaggcaagctg 1015
 QY 970 CTAAGATTACAGTTGATAAGATAGACACATAATTTGTAAGGTTTCAGGAAGTTTCAGAA 1029
 Db 1016 gaaggattgtgatacaagaacacacccagctcgtagatgcaaaagccatagcgatg 1075
 QY 1030 CTATTCTTAACCTATGCTACTGATTAATCGCAATTAGAACAACAACCTCTGACTTTG 1089
 Db 1076 atgttaagacagagtcgagcatgatacaaaccccaaatgcaadtacgacaagcattatg 1135
 QY 1090 ACCGTGAAAAAATFACAAGACGTTTGGCGAAATTTAGCTGGTGTGAGCTGTTATCAAG 1149
 Db 1136 acaaaagaataattgcaagaagaattggctaaactctctggtgaggtggtgctgattaaag 1195

QY 1150 TAGAGCTCCACAGACAGACAGCTTTTAAAGAAATGAACCTTCGATTCAGGATGCTCTAA 1209
 Db 1196 tgggctgctgagtgagtgaaatgaaagagaaaaaacccgggtgtagcgctgtga 1255
 QY 1210 ATGCTACACGCTGACGCGTTTGAAGAGGTATCGTGTGCTGGTGGAAACAGCACTTATTA 1269
 Db 1256 gcgcgactaaagcggcggttgaaagagggcattgtagttggcggtgagcgctctcattc 1315
 QY 1270 CGGTTATTGAAAAAGTAGCAGCTCTTTAGCTTGAAGGCGCATGATGCTGACGCTAACA 1329
 Db 1316 gcgcggtcacaagaagt---gcatttgaattgcagcatgatgaaaaagtggctatgaa 1372
 QY 1330 TTGTGCTTCTGCTCTAGAGAGCGCTGTAGCTCAAAATGCTTTTAAATGCTGGGTACGAAG 1389
 Db 1373 tcatactgctgcgcgcattaaagcccatctagctcaaatcgctatcaacgctggttatgatg 1432
 QY 1390 GCTCCGTAGTATTGCAAGTTGAAAAACAGCCCTGCAGGAACAGGATTTAATCTCCAA 1449
 Db 1433 gcgggtggtggtgagtgaaagtagaaaaacacagagggcatttggtttaacgctagca 1492
 QY 1450 CAGGTGAGTGGGTGTGATATGATTAAACAGCAATCATTTGACCCCTGTCAAAGTAACACGAT 1509
 Db 1493 atggcaagtagtggtatgtttaaagaaggaattattgacctttaaaagttagaaagga 1552
 QY 1510 CAGCGCTTCAAAATGACGCTTCTGTAGTCTTATTTTGAACAACAGAGAGTGTGTTG 1569
 Db 1553 tcgctctacaaaatgggtgttcggttttcaagcgtcttttaaccacagaagccaccgt-- 1610
 QY 1570 CTAAATACCTGACACAGCTTACCCAGCGCCAGCAATGCCAGAGGTTATGATCCAGGAA 1629
 Db 1611 -gcataaatcaagaagaaaaagcactccgcaatgcctgatggtggcgatggcg 1669
 QY 1630 TGATGGTGGGATGGCGG 1648
 Db 1670 gtatggaggtcgtggcg 1688

RESULT 3
 US-08-466-662-7 : Sequence 7, Application US/08466662B
 : Patent No. 6130059
 : GENERAL INFORMATION:
 : APPLICANT: Covacci, Antonello
 : APPLICANT: Bugnoli, Massimo
 : APPLICANT: Telford, John
 : APPLICANT: Macchia, Giovanni
 : APPLICANT: Rappuoli, Rino
 : TITLE OF INVENTION: Helicobacter Pylori Proteins Useful For Vaccines And
 : FILE REFERENCE: CHIR0057
 : CURRENT APPLICATION NUMBER: US/08/466,662B
 : NUMBER OF SEQ ID NOS: 8
 : SOFTWARE: Patent In Ver. 2.1
 : SEQ ID NO 7
 : LENGTH: 1838
 : TYPE: DNA
 : ORGANISM: Helicobacter pylori
 US-08-466-662-7

Query Match 31.8%; Score 528.6; DB 3; Length 1838;
 Best Local Similarity 58.9%; Pred. No. 1.4e-146;
 Matches 966; Conservative 0; Mismatches 664; Indels 9; Gaps 3;

QY 13 ATATGCAAAAGAAATCAATTTTCAGCAGATGCGCGTGTGCCATGGTGGCGGAGTTG 72
 Db 56 aaatggcaaaagaatcaaattttcagatagtcgagaaaaccccttttatttgaagggcgtga 115
 QY 73 ATATGTTAGCAGATACCGTCAAAAGTAAAGCTTGTCTTAAAGCCGCAATTTGTTCTTG 132
 Db 116 ggaactccatgacgctgtcaagtaacacatgggccaagagcgaggaatgtattgatcc 175

:	REGISTRATION NUMBER:	25,146
:	REFERENCE/DOCKET NUMBER:	03495.0137-02000
:	TELECOMMUNICATION INFORMATION:	
:	TELEPHONE:	(202) 408-4000
:	TELEFAX:	(202) 408-4400
:	INFORMATION FOR SEQ ID NO:	28:
:	SEQUENCE CHARACTERISTICS:	
:	LENGTH:	2284 base pairs
:	TYPE:	nucleic acid
:	STRANDEDNESS:	double
:	TOPOLOGY:	linear
:	MOLECULE TYPE:	DNA (genomic)
:	US-08-467-822-28	

	Query Match	29.9%;	Score 496.6;	DB 2;	Length 2284;
	Best Local Similarity	58.1%;	Pred. No. 4.4e-137;	Mismatches 0;	Gaps
	Matches 953;	Conservative			
Qy	13	ATATGGCAAAAGAAATCAAATTTTCAGCAGATGCGCGTGCCTTAAGGGCGGAATGTTCCTTG	72		
Db	504	AAATGGCAAAAGAAATCAAATTTTCAGATAGCGCAAGAACCTTTATTTCGACGGCTAA	563		
Qy	73	ATATGTAGCAGATACCGTCAAAGTAAACCTTGGTCCCTAAGGGCGGAATGTTCCTTG	132		
Db	564	GACAACCTCCATGACCGTGTCAAAGTAAACCATGGGCCAAGAGCGGCAACCGTGTGTATCC	623		
Qy	133	AMAAGCTTTTGGTCTCCCTTAATTAATAAGCGGGTAACCATGTCTAAAGAGATCG	192		
Db	624	AAAAAGCTATGCGGCTCCAAGCATACCAGAAGCGGCTGAGCGTGGCTTAAGAGATG	683		
Qy	193	AATTAGAAGATCATTTTTGAAAACATGGGAGCAAAAATGGTGTCTGAAGTGGCTTCTAAAA	252		
Db	684	AATTAAGTTGCCCGTGGCTTAACATGGGCGCTCAGCTCGTTAAAGAAGATCGAGCAAAA	743		
Qy	253	CCAATGATATTGCTGGTAGGGAGCAGACTACTGCAACAGTTTTTGACACAGCCATGTGTC	312		
Db	744	CCGCTGATGTCGGCGCGGATGGCACGACACAGCAGCCGCTGCTGGCTTATAGCATTTTAA	803		
Qy	313	ATGAAGGACTFAAAAAATGTGACAGCAGGTGCTTAATCCAATTTGGTATCGCTCGAGGCATTG	372		
Db	804	AGAGGGCTTGAGGAATATCAGCGTGGGGCTTAACCCCTATTGAAAGTGAACAGGCATGG	863		
Qy	373	AAACAGCAACAGCAACAGCTGTGTGAAGCCTTGAAAGCCATTTGCTCAACCTGTATCTGGCA	432		
Db	864	ATAAAGCGCTGAAGCGATCATTAATGAGCTTTAAAAAGCGAGCAAAAAGTGGCGGTA	923		
Qy	433	AGGAAGCTATTGCTCAGGTTCGCTGCAGTATCATCAGCTC---TGAAAAAGTTGGAGAGT	489		
Db	924	AGAAGAATAATCACCCAAGTAGGGGACCATTTCTGCAAACTCCGATCACAATATCGGGAAAC	983		
Qy	490	ATATCTCAGAAGCTATTGGAGCGTGTGGGCAACGATGGTGTGATTACCATCGAACAATCTC	549		
Db	984	TCATCGGTGACGCTATGGAAAAAGTGGGTAAAGACGGCGGTATCACCGTTGAAGAAGCTA	1043		
Qy	550	GAGGTATGGAACAGAACTTGAAGTGGTTTGAAGGCATGCAATTTGACCGTGGTTACCTGT	609		
Db	1044	AGGCAATTGAAGATGAATTAGATGTCGTAGAAGGCATGCAATTTTCATAGAGCTACCTCT	1103		
Qy	610	CTCAATACATGGTTCACAGCAATGAAAAAATGGTTGACAGCCTTGAAAAACCATTTATCT	669		
Db	1104	CCCCCTACTTTGTAAACCAACCGCTTGAAAAATGACCGCTCAATTTGGATAACCGTTTACATCC	1163		
Qy	670	TATACAGGATAAAAAGTGTCAACATCCAGACATTTTGCCACTACTCTGAGGAAGTTC	729		
Db	1164	TTTTTACAGGATAAAAAAATCTCTAGCATGAAAGACATTTCTCCCCTACTAGAAAAACCA	1223		
Qy	730	TTAAACCAACCCGCTCCATTTACTCATTTATGACAGATGATGTGGATGGTGAAGCACTTCCAA	789		
Db	1224	TGAAGAGGCCAAACCGCTTTTAATCATCGCTGAAGACATTTGAGGGCGAGCTTTACGA	1283		
Qy	790	CCCTTGTCTTGAACAAGATTCGTGGTACTTTCAATTTGCTTGGTGTCTAAAGCGCCAGAT	849		

;; COUNTRY: USA
;; ZIP: 98121
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/997,080
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Sleath, Janet
;; REGISTRATION NUMBER: 37,007
;; REFERENCE/DOCKET NUMBER: 11000.1007
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 206-269-0565
;; TELEFAX: 206-269-0563
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 159:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1626 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; US-08-997-080-159

Query Match 26.9%; Score 446.8; DB 2; Length 1626;
Best Local Similarity 54.9%; Pred. No. 1.9e-122;

Mismatches 880; Conservative 0; Mismatches 722; Indels 0; Gaps 0;

QY 15 ATGGCAAAAGAAATCAAAATTTTCAGCAGATGCGCGTGTGCGCATGTTGCGCGGAGTTGAT 74
DB 1 ATGGCAAAAGAAATTTTCAGCAGATGCGCGTGTGCGCATGTTGCGCGGAGTTGAT 74
QY 75 ATGTTAGCAGATACCTCAAGCTTGGTCTTAAAGGGCGCAATTTGTTCTTGA 134
DB 61 GCCCTCGCAGACCGCTAAAGGTGAGTTGGCGCGCAAGGGTGGCGCATGTTGAG 120
QY 135 AAAGCTTTTGGTCTCCCTTAATTAATAGCGGGGTAAACCATGCTTAAAGAGATCGAA 194
DB 121 AAGAAGTGGGGCGCCCAACAGATCACCACGATGCTGCTCCATGCCAAGAGATCGAG 180
QY 195 TTAGAGATCATTTTGAACACATGGGAGCAAAATTTGGTCTTGAAGTGGCTTCTAAAC 254
DB 181 CTGGAGGACCGGTACGAGAGATCGCGCTGAGCTGGTCAAGAGGTGCGCAAGAGAC 240
QY 255 AATGATATTGCTGGTATGGGACGACTACTGCAACAGATTTTGACAAAGCCATTTGTCAT 314
DB 241 GAGCAGCTCGCGGCGCAGCGACACCAACCGCCGCTGCTGCTCAGGCTTGGTTCCG 300
QY 315 GAAGGACTAAAATGTGACAGAGGTGCTAATCAATTTGGTATCGTCCGAGGCAATGAA 374
DB 301 GAAGGCTCGCAACGTCGACGCGCGCAACCGCTCGGCTCAAGGCTGGCATCGAG 360
QY 375 ACAGCAACAGCAACAGCTTTGAAGCCTTGAAGCCATTTGCTCAACCTGTATCTGGCAAG 434
DB 361 AAGGCTGTGCGAGGCTGTACCCAGTGGCTGCTGAAGTGGCGGCAAGGAGTGGAGACCAAG 420
QY 435 GAAGTATTGCTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 494
DB 421 GAGCAGATTTCTGCCACCGCGCGGATTTCCGCGCGGCGACACCCAGATCGCGAGCTATC 480
QY 495 TCAGAGCTATGAGCGGTGGGCAACGATGCTGATTAACATCAAGAAATCTCGAGT 554
DB 481 GCCGAGGCGCATGGAAGGTGGGCAACGAGGCTGCTATCAGCGTGGAGAGTGAACAC 540
QY 555 ATGGAACAGCAATTCAGTGGTATGAAGGCAATTTGACCTGGTGTACCTGTCTCA 614

DB 541 TTCGGCCTCGAGCTCGAGCTACCGAGGGTATCGCGTTGCACAAGGGCTACATCTCGGT 600
QY 615 TACATGGTCACAGACAATGAAAAATGGTGTGACAGCTTTGAAAAACCAATTTATCTTAATC 674
DB 601 TACTTGTGACCGACCGCGAGCGCGTCTTGAGGATCCCTACATCTCTGCTGCTG 660
QY 675 ACGGATAAAAAAGTGTCAACATCCAAAGACATTTTCCACTACTTCTGAGGAAGTCTTAA 734
DB 661 GTGAGCTCCAAGGTGTGACCGCTCAAGGATCTGCTCCGCTGTGGAAGGTCTATCCAG 720
QY 735 ACCAACGCTCCATTTACTTATTTGAGATGATGTGGATGGTGAACACTTCCAAACCTT 794
DB 721 GCGGCAAGCGCTGCTGATCATCGCGGAGGCTGAGGGCGAGGCTGTCCACGCTG 780
QY 795 GTCTTGAACAAGATTCGTGTACTTTCAATGTGGTGTCTCAAGCGCCAGGATTTGGT 854
DB 781 GTGGTCAACAAGATCCGCGCACCTTCAAGTCCGCTCAAGGCTCGGGCTTCGGT 840
QY 855 GATCGTCTAAAGCTATGCTTGAAGACATTTGCTATTTGACAGGTGTGACAGTATACA 914
DB 841 GACCGCGCAAGCGCATGCTGCGAGGACATGGCCATCTCACCGGTGTGAGGTCTCAGC 900
QY 915 GAGGATCTAGGACTTGAATTAAGAATGCTACATACAGACCCCTTGGACAGGCTGCTAAG 974
DB 901 GAAAGAGTCCGGCTGTCTGAGAGCCGCGGAGCTCTCGCTGCTGGCGCAGGCCCGCAAG 960
QY 975 ATTACAGTTGATAAGATAGCAGTAATTTGTAAGGTTTCAGAGATTCAGAAAGCTATT 1034
DB 961 GTGCTGCTCACCAGGACGAGACCCACCATGCTGCGAGGGTTCGGGCGATTCGGATGCCATC 1020
QY 1035 GCTAACGCTATTGCACTGATTAATCGCAATTAGAAACAACTTCTGACTTTTGACCGT 1094
DB 1021 GCGGCGCGGTGGTCTCAGATCCGCGCGAGATCGAGAACAGGCTCTCCAGTACGACCGC 1080
QY 1095 GAAAACTACAGAGAGCTTTGGGGAATTTAGCTGGTGTGAGTGTATCAAAAGTAGGA 1154
DB 1081 GAGAAGCTCGAGAGCGCTGCGCAAGCTGGCGCGGCTGTTGCGGTGATCAAGGCGGA 1140
QY 1155 GCTCAACAGAGACAGCTTTTAAAGAAATGAACTTCGCAATTCAGGATGCTCTAAATGCT 1214
DB 1141 GCTGCCACCGAGGTGAGCTCAAGGAGCGCAGCAGCCGATCGAGGACCGCTCCGCAAC 1200
QY 1215 ACAGTGTGACGCGTTGAAGAAGTATGCTTGTGTTGGTGGTGAACAGCACTTATACGGTT 1274
DB 1201 GCGAAGGCTGCGGTGAGAGGCGCATGCTGCGCGGTGGCGGTGCTGCTGCTGCTGCTG 1260
QY 1275 ATTGAANAAGTAGCAGCTCTTGAAGTGTGAGGCGGATGCTACTGAGAGTACATTTG 1334
DB 1261 GCTCTGCGCTGGAGACCTCGCGCTGAGCGGCGAGGCGCACCGGTGCCAACATCGTC 1320
QY 1335 CTTCGTGCTCTAGAAGAGCTGTAGCTCAATTTGCTTTAAATGCTGGGTACGAAGGCTCC 1394
DB 1321 CCGGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
QY 1395 GTAGTTATTGACAAGTTGAATAACAGCCCTGCAAGAACAGGATTTAATGCTGCAACAGT 1454
DB 1381 GTGCTTGGCGAAGGTGTCCAACCTGCGCGGTGCGCGGTGCTGCTGCTGCTGCTGCTG 1440
QY 1455 GAGTGGTGTATGATTAATAACAGGAATCATTTGACCTGTCAAAAGTACAGATCAGG 1514
DB 1441 GAGTACGAGGAGCTGCTCAAGCGCGCGCTGCGCGGAGGCTCAAGGCTCAGCGCTCG 1500
QY 1515 CTTCAAATGCAAGCTTCTGCTAGTCTTATTTTGAACAACAGAGAGCTTGTGCTAAT 1574
DB 1501 CTGCAAGACCGCGGTCTCATCGCGCTCTGTTCTTCAACCGAGGCGCTGCTGCGCGAC 1560
QY 1575 AAACCTGAACAGCTTACGCCAGCGCAGCAATTCGCAAGAGT 1616
DB 1561 AAGCGGAGAGGCGTCCGCAACCGCGGCGGCGACCGCGGT 1602

RESULT 6
US-08-997-362-159

QY	315	GAAGGACTAAAAATGTGACAGCAGGTGCTTAATCCAAATTTGGTATCCGTGAGGCATTGAA	374
DB	301	GAAGSCTGGCACAACGTCGACGCGCGGCCAACCGCTCGGCCTCAAGCGTGGCATCGAG	360
QY	375	ACAGCAACAGACAAGCTGTTGAAGCCTTTGAAGCCATTGGAAGCCATTGCTCAACCTGTATCTCTGGGCAAG	434
DB	361	AAGGCTGTCGAGGCTGTCAACCCAGTCGCTCAAGTTCGGCCAAAGAGGTGCGAGACCAAG	420
QY	435	GAAGCTATTGCTCAGGTCGCTGCAGTATCATCACGCTCTGAAAAGATTGGAGAGTATATC	494
DB	421	GAGCAGATTCTGCCACCGCGGCCATTTCGGCGGGCGACACCAGATCGGCGAGCTCATC	480
QY	495	TCAGAAGCTATGGAGCGTGTGGGCAACGATGTGTGATTACCATCGAAGAAATCTCGAGGT	554
DB	481	GCCGAGGCCATGACACAGGTCGCAACGAGGCTGTCATCACCGTCGAGGATCGACACC	540
QY	555	ATGGAACACAGAACTTGAAGTGGTTGAAGGCATGCAATTTGACCGTGGTGTACCTGTCTCAA	614
DB	541	TTCGGCCTGCAGCTCGAGCTCACCGAGGGTATGCGCTTCGACAAGGGCTACATCTCGGGT	600
QY	615	TACATGCTCAGACACATGAAAAAATGGTTGCAGACCTTTGAAAACCCATTTATCTTAATC	674
DB	601	TACTTCTGTGACCGACCGCGGAGCGCCAGGAGCGCTCTCGAGGATCCCTACATCTCGTGTG	660
QY	675	ACGGATAAAAAGTGTCAAACATCCAAACATTTTGGCCACTACTTTGAGGAAGTTCCTAAA	734
DB	661	GTCAGCTCCAGGTGTGCGACCGCTCAGGATCTGCTCCCGCTCTCGAAGAGGTATCCAG	720
QY	735	ACCAACCGTCCATTACTCATTTATTGCAGATGATGGATGGTGAAGCACTTTCCAAACCCTT	794
DB	721	GCCGGCAAGCGCTGCTGATCATCTCGCAGGAGAGTCTGAGGGCGAGGCCCTGTCCACGCTG	780
QY	795	GTCATTGACAGATTCGTGTACTTTCATGTGTGTTGCTGTCAAAGCGCCAGGATTTGGT	854
DB	781	GTGTCTAACAGATTCGCGCGCACCTTTCAAGTCGCTGCGCTCAAGGCTCCGCGGCTCGGT	840
QY	855	GATCGTCGTAAGCTATGCTTTGAAGACATTTGCTATCTTCACAGGTGGTACAGTGATTACA	914
DB	841	GACCGCGCAAGCGATGCTGCAGGACATGGCCATCTCTCACGGTGGTCAGGTGTCAGC	900
QY	915	GAGGATCTAGGACTTGAAATTTAAAGATGCTACAATGACAGCCCTTTGGACAGCGTCTAAG	974
DB	901	GAAGAGTCGGGCTGTCCCTGGAGACGCGCAGCTCTCCGCTGCTGGCGCAGCCCGCAAG	960
QY	975	ATTACAGTTGATAAAGATACACAGTAATTGTTGAAGGTTTCAAGAGTTTCAAGAGTCTATT	1034
DB	961	GTCCTCTCTACCAAGGACGAGACCCATCTGCTGAGGGCTCGGGCGATTCCGATGCCATC	1020
QY	1035	GCTAACCGTATTGCATGATTAAATCGCAATTAGAAACAACAATCTTGACTTTTGACCGT	1094
DB	1021	GCGGCGGGTGGCTCAGATCCGCGCGAGATCGAGAACAGCGACTCCGCTACGACCGC	1080
QY	1095	GAAGAACTACAAGAACGTTTGGCGAAATTAGCTGGTGGTGTAGCTTGTATCAAGATGAGA	1154
DB	1081	GAGAAGCTGACGAGCGCTTGGCCAAAGCTGGCGGGGCTGTTGCGGTGATCAAGSCCGA	1140
QY	1155	GCTCCACAGACAGACGTTTAAAGAAATGAACCTTCGCAATTCAGGATGCTCTTAATGCT	1214
DB	1141	GCTGCCACCGAGTGGAGCTTCAAGGAGCCGACAGCACCGCATCGAGGAGCGGCTCCGCAAC	1200
QY	1215	ACAGTCGACCGTTGAAGAAGTATCTGTTGTTGGTGGTGGAAACAGCACTTATTACGCTT	1274
DB	1201	GCGAAGCTCCGCTCGAAGAGGCGATCTGTCGCGGCTGGCGGGCTGGCTCTGTCAGTCTG	1260
QY	1275	ATTGAAAAGTAGCAGCTTTGAGCTTTGAGGCGGATGATGCTTACTGGACGTAACATTTGTG	1334
DB	1261	GCTCTCGGCTGGACGACCTCGGCCCTGACGGGGGAGAGGCCACCGGTCCCAACATCGTC	1320
QY	1335	CTTCGTGCTCTAGAAGAGCCCTGTACGTCACAAATTTGCTTTAAATGCTTGGGTACGAAGGCTCC	1394
DB	1321	CGCGTGGCGCTGTGCGCTCGCTCAAGCAGATCGCCTTCAACGGCGGCTCGGAGCCCGGC	1380

QY 915 GAGATCTAGACTTGAATTAAGAATGCTACAAATGACACCCCTTGGACAGGCTGCTAAG 974
Db 901 GAAAGAGTCGGGTGCTCCCTGGAGACCGCGAGCTCTCGTGTGGCCAGCGCGCAAG 960
QY 975 ATTACAGTTGATAAAGATACACAGTAATTTGAAGTTTCAGGAAGTTTCAGAAAGCTATT 1034
Db 961 GTCTGCTCACCAGGACGAGACCACTCGTCAGGGCTCGGGCGGATCCGATGCGATC 1020
QY 1035 GCTAACCGTATTGACTGATTAATGCAATTAAGAACAACAACTTCTGACTTTGACCGT 1094
Db 1021 GCCGCGGGTGTCTCAGATCCGCGGAGATCGAGACAGGACACTCCGACTACGACCGC 1080
QY 1095 GAAAACTACAAGACGTTGGCGAAATTAAGTGGTGGTGTAGCTGTTATCAAAAGTAGGA 1154
Db 1081 GAGAAGCTGACGAGGCGCTGGCGCAAGCTGGCGGGGGTGTTCGGGTGATCAAGGCGGA 1140
QY 1155 GCTCCACAGACAGAGCTTTAAAGAAATGAAACTTCGCAATTCAGGATGCTCTAAATGCT 1214
Db 1141 GCTGCCACCGAGGTGAGCTCAAGGAGCGCAACCGCATCGAGGACCGCGTCCGCAAC 1200
QY 1215 ACAGTGCAGCCGTTGAAGAAGTATCGTTGCTGGTGGTGGAAACAGCACTTATTACGGTT 1274
Db 1201 GCGAAGGCTCGCTCGAAGAGGCGATCGTCGCGCGTGGCGGCTGCTGCTGCACTG 1260
QY 1275 ATTGAAAAAGTACAGCTCTTGAAGTGGAGGCGGATGATGCTACTGGACGTAACATTTG 1334
Db 1261 GCTCTGCGGTGACGACCTCGGCTGAGCGGGGAGGAGCCACCGGTGCAACATCGTC 1320
QY 1335 CTTGCTGCTCTAGAACAGCTGTACGCTCAAAATGCTTTAAATGCTGGTACGAAGCTCC 1394
Db 1321 CGGCTGGCGTGTGCGCTCGCTCAAGCAGATCGCCTCAACGGCGCGCTGGAGCCGCG 1380
QY 1395 GTAGTTATTGACAAAGTTGAAAAAGCCCTGCGAGGAACAGGATTTAATGTCGAACAGGT 1454
Db 1381 GTGTTGCGGAGAAGGTGTCACACCTGCCGCGGTCACGGCTCAACGCGGACCGGT 1440
QY 1455 GAGTGGTGTATGATTAACAGAGAAATCAATGACCCCTGTCAAAATCAACAGATCAGCG 1514
Db 1441 GAGTACGAGGACGTGTCGAAGGCGCGCTGCGGAGCCGCTGAAGTTCACCCGCTCGGG 1500
QY 1515 CTTCAAAATGACGCTCTGTAGTCTGCTTATTTTTCACACACAGCAAGCAGTTGCTAAT 1574
Db 1501 CTCACAAACGCGGCTCCATCGCGGCTCTGTTCTCACCACCGAGCGCGTGTGCGCGAC 1560
QY 1575 AAACCTGA 1582
Db 1561 AAGCCGGA 1568

RESULT 9

US-08-997-362-113
; Sequence 113, Application US/08997362
; Patent No. 5985287
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Hiyama, Jun
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Scott, Linda
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
; TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/997,362
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970
FILING DATE: June 12, 1997
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705,347
FILING DATE: August 29, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002c2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 1569 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-08-997-362-113

Query Match 26.8%; Score 444.8; DB 2; Length 1569;

Best Local Similarity 55.2%; Pred. No. 7.5e-122;

Matches 866; Conservative 0; Mismatches 702; Indels 0; Gaps 0;

QY 15 ATGGCAAAAGAATCAAAATTTTCAGCAGATGCGCGTCTGCCATGTCGCGGAGTTGAT 74
Db 1 ATGGCCAAGACAAATTCGGTATGACGAAGAGCCCGCGCTCAGCGGGCCCTCAAC 60
QY 75 ATGTTAGCAGATACCTCAAAATTAAGCTTGGTCTTAAAGGGCGCAATGTTGTTCTTGA 134
Db 61 GCCTCGCAGACCGCTAAAGGTGAGTGGGCCCGCAAGGTCGCAACGTCGTCGTCGAG 120
QY 135 AAGCTTTTGGTCTCCCTTAATTAATTAAGTACGAGGCTAACCATTGCTAAAGAGATCGAA 194
Db 121 AAGAAGTGGGCGCGCCACGATCAACACGATGTTGTCTCCATCCCAAGGAGATCGAG 180
QY 195 TTAGAAGATCATTTTGAACACATGGGACAAATTCGTTGCTGAAGTGCCTCTAAACCC 254
Db 181 CTGAGGACCCGTACGAGAAGATCGCGCTGAGCTGTTCAAGAGGTCGCCAAGAGACC 240
QY 255 AATGATATTGCTGGTATGGGACGACTACTGCAACAGTTTTCACACAGCCATTTGTTCA 314
Db 241 GAGCAGCTCGCGGCGAGCGCACCCACCGCCAGCTGCTCGCTCAGGCTCTGGTTCGC 300
QY 315 GAAGGACTAAAATATGTCAGCAGGTGCTAATTAATTAATTAATTAATTAATTAATTA 374
Db 301 GAAGGCTCGCAACGTCGACGCGCGCAACCGCTCGGCTCAAGCTGCGCATCGAG 360
QY 375 ACAGCAACAGCAACAGCTGTTGAAGCTTGAAGCCATTGCTCAACCTGTATCTGCAAG 434
Db 361 AAGGCTGTCAGGCTGTCACCCAGTGCCTGTAAGTTCGCGCAAGGAGTTCGAGACCAAG 420
QY 435 GAAGTATTGCTCAGTCTGCTCAGTATCATCATCATCATCATCATCATCATCATCATCAT 494
Db 421 GAGCAGATTTCTGCCACCGCGCGGATTTCCGCGGCGGACACCCAGCATCGCGGCTCATC 480
QY 495 TCAGAGCTATGAGCGGTGTCGCAACGATGTTGTTGATTAATTAATTAATTAATTAATTA 554
Db 481 GCCGAGGCTGCAAGGTCGCAACGATGTCGCAACGATGTCATCACCGTTCGAGGAGTCAAC 540
QY 555 ATGGAACAGAACTTGAAGTGGTTGAAGCATGCAATTTGACCGTGGTGTACCTGCTCA 614
Db 541 TTCGCGCTGACGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCT 600

[illegible]

RESULT 10

RESULT TO
US-09-095-855-113

US-09-093-833-113
; Sequence 113, Application US/09095855

; Sequence 113, Appl.
; Patent No. 6160093

GENERAL INFORMATION:
APPLICANT: Tan, Paul
APPLICANT: Visser, Elizabeth
APPLICANT: Skinner, Margot
APPLICANT: Prestidge, Ross
TITLE OF INVENTION: Compounds and Methods for
TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
NUMBER OF SEQUENCES: 208
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/095,855
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/705,347
FILING DATE: 29-AUG-1996
APPLICATION NUMBER: 08/873,970
FILING DATE: 12-JUN-1997
APPLICATION NUMBER: 08/997,362
FILING DATE: 23-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002c3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 1569 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
PS-09-095-855-113

[illegible]

Db 1 ATGCCGAAGATTCTGAAGTTCCAGCAGGACGCCCTCGCGCCCTTGAGCGCGCGTGAAC 60
QY 75 ATGTTAGAGATACCGTCAAGATTAACGTTGGTCTCTAAAGGCGCAATGTTGTTCTTGA 134
Db 61 CAGCTGGCCAGACACCGTCAAGTGACCATCGGCCCAAGGCGCGCAACGCTGTCATCGAC 120
QY 135 AAAGCTTTTGGTCTCTCCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 194
Db 121 AAGAAGTTGGCGCCCGCCAGCATCACCAGCAGCGCGTCAACATCGCCCGTGAGGTGAG 180
QY 195 TTAGAAGATCATTTTGAAGCATGGGAGCAAAATTTGGTCTGAAGTGGCTTCTTAAAC 254
Db 181 TGCAGCAGCCGTTACGAGAACCTCGCGCCAGCTCGTCAAGGAGGTGGCCACCAAGACC 240
QY 255 AATGATATTGCTGTGATGGAGCAGCTACTGCAACAGTTTGGACACAGCCATTTGTTTAT 314
Db 241 AACACATATCGCGGTGACGCGCACCAACCGCGACCTGCTGCGCCAGCGGCTGTCGCG 300
QY 315 GAAGGACTAAAAATGTACAGCAGGTGCTTAATCCAAATTTGGTATCCGTGAGGCAATTGAA 374
Db 301 GAGGCGCTGGCAACGTCGCGCGCGCGCGCTCCCGCGCGCTTGAAGAGGCGATCGAC 360
QY 375 ACAGCAACAGCAGCATGTTGAAGCCCTTGAAGCCATTGCTCAACCTGTATCTGGCAAG 434
Db 361 GCGCCGCTCGCGCGCTGTCGCGCGAGCTGCTCAGACCGCGCGCGCGATCGACGACAAG 420
QY 435 GAAGCTATTGCTCAGGTCGCTGAGTATCATCATCAGCTCTGAAAGAGTTGGAGAGTATC 494
Db 421 TCCGACATCGCGCGCTGCGCGCTCTCCGCGCAGGACAGCAGTGGTGGCGGAGCTCATC 480
QY 495 TCAGAACTATGGAGCGTGTGGGCAACGATGGTGTGATTACCATCGAAGAAATCTGAGGT 554
Db 481 GCGGAGCGATGGACAAGGTGCGCGAGGAGCGTGTGTCATCCCGTCGAGGAGTCCAAAC 540
QY 555 ATGGAACACAGATTTGAAGTGGTTGAAGGATGCAATTTGACCGTGTACTGTCTCAA 614
Db 541 TTCGGTCTCACTGGACTTCCAGGAGGATGCGCTTCACAGGCGTACCTGTCCCG 600
QY 615 TACATGTCTCAGACAAATGAANAATGGTTGACACCTTGAACCCCATTTATCTTAATC 674
Db 601 TACATGTGACCGACAGGAGGTGAGGCGCTCTCGACGACCGGTACATCTGATC 660
QY 675 ACGGATAAAAGTGTCAACATCCAGACATTTTCCACTACTTGTAGGAAAGTTCTTAA 734
Db 661 CACGAGGCAAGATTCGTTGATCAGGAGCTGCTGCGCTGTGGAAGGTGATCCAG 720
QY 735 AC-----CAACCGTCCATTAATTTGACATGATGTTGATGTTGAAGCACTTCCA 788
Db 721 GCGGTCGCTTCAAGCGCTGCTGATCATCGCCGAGGACGTCGAGGCGGCGCTGTG 780
QY 789 ACCCTGCTTGAACAGATTCGTTGTTTCAATGTTGTTGTTCAAGCGCCAGGA 848
Db 781 ACCCTGTTGTTCAACAGATTCGCGGCGACGTTCAACGCGCTGCGCGTCAAGGCGCCG 840
QY 849 TTTGCTGATCGTCAAGATGCTTGAACAGATTCGTTGATGTTGAGGAGTGTGACAGT 908
Db 841 TTCGTTGACCGCCCAAGGAGGATGCTCGGCGACATGCGCCCTCACCGCTGCGCCGTC 900
QY 909 ATTACAGGATCTAGGATTTGAATTAAGATGCTACATGACAGCCCTTGGACAGCT 968
Db 901 ATCCCCAGGAGTTCGCTCAAGCTGACCGCGCGCTGAGCGTGTGGGCGCCG 960
QY 969 GCTAAGATTACAGTTGATAAGATPAGCAGATTAATTTGAAGTTTCAAGGAGTTCA 1028
Db 961 CGCGCGTCAACCGTCAACAGGAGGACGACGACATCGTGGAGCGCGCGCAACCGCGAG 1020
QY 1029 GCTATTCTTACCGTATTGACATGATTAATTCGCAATTTAGAACACAACTTCTGACTTT 1088
Db 1021 GAGTTCAGGCGCGCTGCGCCAGATCAAGCGCGGAGATCGATCGACCGACTCGGACTGG 1080
QY 1089 GACCGTGAAGAACTACAGAGCGTTTGGCGGAATTAAGTGTGTTGTTATCAAA 1148
Db 1081 GACCGGAGAGGCTCCAGGAGCGCTCCCAAGCTGGCGCGCGCTGCTGCGTATCCG 1140

QY 1149 GTAGAGCTCAACAGACAGACAGCTTTAAAGAAATGAACATTCGCATTTAGAGATGCTCTA 1208
Db 1141 GTCGCGCGCGCCAGCTGAGCTGAAGGAGCGCAACGACCTCTCTGAGGAGCGCCATC 1200
QY 1209 AATCTACAGCTGAGCGCTTGAAGAGGTATCTGCTGCTGTTGGACAGCACTTAT- 1267
Db 1201 TCCGCGACCGCGCGCGCTGAGGAGGCGATCTCTCCGCTGCTGCTGCTGCTGCTGCT 1260
QY 1268 --TACGCTTATTGAAAGTAGCAGCTTCTGAGCTTGAAGCGGATGATGCTACTTGGAGCT 1325
Db 1261 CACGCGCTCAAGTCTCTGAGCAACCTCGCGCGCACCGCGGAGCGCCACCGTGTGTC 1320
QY 1326 AACATTTGCTTCTGCTCTGAGAGAGCTGTACGTCAAAATGCTTTAAATGCTGGGTAC 1385
Db 1321 CGGTCGCTCGCGCGCGCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
QY 1386 GAAGGCTCCGTAGTATTGACAAGTTGAAAGAGCTTGAAGAGCTTGAAGAGCTTAAATGCT 1445
Db 1381 GAGGCTACGCTCATCACCAAGGTGGCGAGCTCGAAGGCGCGAGGCTTCAACCG 1440
QY 1446 GCAACAGGTGAGTGGTGTATGATTAATAACAGGAATCATTTGACCTGTCAAGTAACA 1505
Db 1441 GCCACCGGAGTACGCGGACCTGCTCAAGCGCGCGCTCATCGACCGGTCAGGTCACC 1500
QY 1506 CGATCAGCGCTTCAAAATGACAGTCTCTGTAGTCTTATTTTGAACAAAGAGCACTT 1565
Db 1501 CGGTCGCGCTGAGAGCGCGCTTCCATCGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
QY 1566 GTTCTTAATAACCTGAACAGCTACCGCAGCGCAG 1602
Db 1561 GTCGTCGAGAAGCGCGCGGAGGAGGAGGCGCGGAGGCG 1597

RESULT 12
US-09-031-606-10
; Sequence 10, Application US/09031606
; Patent No. 6153404
; GENERAL INFORMATION:
; APPLICANT: MAZODIER, Philippe
; APPLICANT: GUGLIEMI, Gerard
; TITLE OF INVENTION: REGULATORY NUCLEOTIDE SEQUENCE OF THE
; TITLE OF INVENTION: INITIATION OF TRANSCRIPTION
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: George Mason Bldg., Washington & Prince Sts.
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,606
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/050,313
; FILING DATE: 10-MAY-1993
; APPLICATION NUMBER: -FR 9011186
; FILING DATE: 10-SEP-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feury, Sharon E
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 010830-035
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 1620 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURES:

NAME/KEY: CDS
LOCATION: 1..1620
US-09-031-606-10

Query Match 25.1%; Score 416.6; DB 3; Length 1620;

Best Local Similarity 54.7%; Pred. No. 1.7e-113;

Matches 874; Conservative 0; Mismatches 714; Indels 9; Gaps 2;

QY 15 ATGGCAAAAGAAATCAATTTTTCAGCAGATGCGCGTGTCTGCCATGTGTCGCGGAGTGTGAT 74
DB 1 ATGGCGAAGATTCTGAAGTTCGACGAGGAGCGCCGTCGCGCCCTTGAGCGCGCGGTGAAC 60
QY 75 ATGTTAGCAGATACCGTCAAAAGTAAAGCTTGGTCTTAAAGGGCGCAATCTTGTCTTGA 134
DB 61 CAGCTGGCGCACACGCTCAAGGTGACCATCGGCCCAAGGGCGCAACGTCGTTCATCGAC 120
QY 135 AAGCTTTTGGTCTCCCTTAATTAATAGTACGGGGTAAACATTGCTTAAAGAGATCGAA 194
DB 121 AAGAAGTTCGGGCCCCCGACCATCAACAGAGCGGCTCACCATCGCCCGTGAAGTCGAG 180
QY 195 TTAGAAGATCAATTTGAAGAACATGGGAGCAAAATTTGGTCTCTCAAGTGGCTTCTAAACC 254
DB 181 TGGAGCAGCCGTACGAGAACTTCGGCGCCAGCTCGTCAAGGAGTGGCGACCAAGACC 240
QY 255 AATGATTAATTTGGTGTGATGGGAGCTACTGCAACAGATTTTGAACAAGCCATTTGTTAT 314
DB 241 AAGCATATCGCGGTGACGCGACCAACACCGCGACGCTGTCGCCACGCGTGTGTCGCC 300
QY 315 GAAGGACTAAATAATGTGACAGAGTGTCTAATCAATTTGGTATCGTTCGAGGCAATTGAA 374
DB 301 GAGGGCTTCGCAACGTCGCGCGCGCTCCCGCGCGCTCAAGAGGGCATCGAC 360
QY 375 ACAGCAACAGCAACAGCTGTTCAAGCCTTGAAGCATTTGCTCAACCTGTATCTGCGAAG 434
DB 361 GCGCGCTGCGCGCGCTTCGCGCGAGCTGCTCGACACCGCGCGCGATCGACGACAAG 420
QY 435 GAAGCTAATTTGCTCAGTGTGCTCAGTATCATCACGCTCTTGAATAAGTTGGAGAGTATATC 494
DB 421 TCCGACATCGCGCGCTGCGCGCTCTCCGCGAGGACAAGCAGTGTGCGGAGCTCATC 480
QY 495 TCAGAACTATGAGGGTGTGGGCAACGATGTTGATTAACCATCGAAGATCTCGAGGT 554
DB 481 GCCGAGGCGATGGACAAGGTTCGGAGGAGCGGTGTCTATCACCGTGGAGGATCCAAACAC 540
QY 555 ATGGAACAGAACTTCAAGTGTGTTGAAGCATGCAATTTGACCGTGGTTACCTGTCTCAA 614
DB 541 TTCGGTGTGACCTTGACTTCAACGAGGCGATGGCGTTCGCAAGGGCTTACCTGTCCCGG 600
QY 615 TACATGGTTCACAGCAATGAAAAATGGTTGTCAGACCTTTGAAAAACCCATTTATCTTAATC 674
DB 601 TACATGGTTCACGACGAGGAGCTATGAGGCGCTTCTCGACGACCGGTACATCTGATC 660
QY 675 ACGGATAAATAAGTGTCAACATCAAGACATTTTGGCACTACTTGAAGAGTCTTTAA 734
DB 661 CACGAGGCAAGATGGTTTCGATCGAGGACCTTGTGCGGCTGCTGGAGAAGTCTATCCAG 720
QY 735 AC-----CAACCGTCCATTACTCATTAATTCAGATGATGTGATGCTCAACACTTCCA 788
DB 721 GCGGTGGTCTCAAGCGCTGCTGATCATFCGCGGAGGAGTGTGAGGGGAGGCCCTGTGCG 780
QY 789 ACCCTTGTCTTGAACAAGATTCGTGCTACTTTCAATTTGGTTGCTGTCAAAAGCGCAGGA 848
DB 781 ACCCTTGGTGTCAACAAGATTCGCGGACGCTTCAACGCGCTCGCGTCAAGCGCGCCGCG 840
QY 849 TTTGGTGTGTCGTTAAGCTAATGCTTGAAGACATTTGCTATCTTGACAGGTGGTACAGTG 908

DB 841 TTCCGTGACCGCGCAAGCGGATGCTCGCGCAGATGCCACCTCCCGGTGCCACCGTC 900
QY 909 ATTACAGAGGATCTAGGACTTGAATTAAGAAGATGCTACAATGACAGCCCTTGGACAGGCT 968
DB 901 ATCGCGGAGGAGGTGCGCGCTCAAGCTCGACAGCGCGGTCTGGACGCTGTGGCACCAGCC 960
QY 969 GCTAAGATTACAGTTGATAAAGATAGCACAGTAATTTGTAAGTTTCAGGAAGTTTCAGAA 1028
DB 961 CGCGCGTCAACGCTCACCAAGGAGGACACGACCATCTGAGCGCGCGGCAACGCGGAG 1020
QY 1029 GCTATTGCTAACCGTATTCGACTGATTAATCGCAATTAAGAACAACACTTCTGACTTT 1088
DB 1021 GACGTCCAGGGCGCGCTCGCCAGATCAAGGGCGAGATCGAGTCGACCGACTCGGACTGG 1080
QY 1089 GACCGTGAATAACTACAACAAGCTTTGGCGAAATTAAGTTCGTTGGTGTAGCTGTTATCAA 1148
DB 1081 GACCGGAGAAAGTCCAGGAGCGCTTCGCCAAGTGGCGGCGGCGTCTGCGGTATCCGC 1140
QY 1149 GTAGGAGCTCCCAACAGAGACAGCTTTAAAGAAATTAAGAACTTCGCATTCGAGATGCTTA 1208
DB 1141 GTCGCGCGCGCACCGAGCTCGAGCTGAAGGAGCGCAACGCTCTGGAGGACGCCATC 1200
QY 1209 AATGCTACAGCTGACGCGCTTGAAGAAGTATCGTTGCTGGTGGTGGAAACAGCACTTAT- 1267
DB 1201 TCCGCGACCGCGCGCGGTTCGAGGAGGAGCATCTCTCGGTGGTGGTCCGCGCTGGTC 1260
QY 1268 --TAGGTTATTGAAAAAGTACAGCTCTTGAAGTTGAGGCGGATGATCTACTGACCT 1325
DB 1261 CAGCGCGTCAAGTCTTGACGACACACCTTCGCGCGGACCGCGGACGAGGCCACCGGTGC 1320
QY 1326 AACATTGTCTTCGTCTCTAGAAGAGCGCTGTACGTCATAATTTGCTTTAATGCTGGGTAC 1385
DB 1321 GCGGTGCTCGCGCGCGCGCTCGAGCGCTGCGGTGATCGCGGAGAACCGCGGCCCTC 1380
QY 1386 GAAGGTCGTGAGTTATTGACAAGTTGAAAAACAGCCCTGACGAGAACAGGATTTAATGCT 1445
DB 1381 GAGGCTAGTCTATCACCACCAAGGTGGCGAGCTCGACAAGGGCGGAGGCTTCAACGGG 1440
QY 1446 GCAACAGGTGAGTGGTTGATATGATTAACACAGAACTTACGCTCTCAAGTAAACA 1505
DB 1441 GCCACGGGAGTACGGCACTGTCAGGCGCGGTCTGACCCCGGTCAAGGTACAC 1500
QY 1506 CGATCAGCGCTTCAAAATCAGCTTCTGTAGTAGTCTTTATTTGACAAACAGAGCAGTT 1565
DB 1501 GCGTCCGCTTGGAGAACGCGCGCTCCATCGCTCCCTCTCTGAGACCGAGACCCCTG 1560
QY 1566 GTTGTATATAACCTGAACCACTGAGCGCGGAGGAGGCGCGGAG 1602
DB 1561 GTCGTGAGAGCGCGCGGAGGAGGAGGCGCGGAGGCGG 1597

RESULT 13

US-08-461-775-11

; Sequence 11, Application US/08461775

; Patent No. 5858773

; GENERAL INFORMATION:

; APPLICANT: MAZODIER, Philippe

; APPLICANT: GUGLEMI, Gerard

; TITLE OF INVENTION: REGULATORY NUCLEOTIDE SEQUENCE OF THE

; TITLE OF INVENTION: INITIATION OF TRANSCRIPTION

; NUMBER OF SEQUENCES: 15

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Burns, Doane, Swecker & Mathis

; STREET: George Mason Bldg., Washington & Prince Sts.

; CITY: Alexandria

; STATE: Virginia

; COUNTRY: United States

; ZIP: 22313-1404

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

— — — — —

APPLICANT: MAZODIER, Philippe

Db 2405 CCTGGTCGTCGAGAAGCCGGCCGAGGAGGCCCGAGGCCG 2445

APPLICANT: GUGLIEMI, Gerard
TITLE OF INVENTION: REGULATORY NUCLEOTIDE SEQUENCE OF THE
INITIATION OF TRANSCRIPTION
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031,606
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/050,313
FILING DATE: 10-MAY-1993
APPLICATION NUMBER: FR 9011186
FILING DATE: 10-SEP-1990
ATTORNEY/AGENT INFORMATION:
NAME: Crane-Feury, Sharon E
REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 010830-035
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)-836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 2668 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-031-606-11

Query Match 24.7%; Score 411; DB 3; Length 2668;
Best Local Similarity 54.5%; Pred. No. 9.9e-112;
Matches 872; Conservative 0; Mismatches 720; Indels 9; Gaps 2;
QY 11 TCATATGCGAAAGAAATCAAAATTTTCAGCAGATCGCGTGTGCTGCCATGGTGGCGCGGAGT 70
DB 845 TCCCATGGCGAAGATTCGAAGTTCGACGAGGACGCCCGTTCGCCCTTGACGCGCGGT 904
QY 71 TGATATGTTAGCAGATACCGTCAAAAGTAACGCTTGGTCTTAAAGGCGCAATGTTGTCT 130
DB 905 GAACAGCTGGCGGACACCGTCAAGGTGACCATCGGCCCAAGGCGCGCAAGCTGTCAT 964
QY 131 TGAAGAACTTTTGGTCTTCCCTTAATTAATTAAGTACGCGGGTAACCATGCTTAAGAGAT 190
DB 965 CGACAAGAAAGTTGCGCGCCCGACCATCAACCAACGACGCGCTCACCATCGCCCGGTGAGT 1024
QY 191 CGAATTAGAAGATATTGTTGAAACATGGAGCAAAATTTGTTGCTGAAGTGGCTTTCAA 250
DB 1025 CGAGTGCACGACCGGTACGAGAACCTCGGCGCCCGAGCTGTCAGAGGAGTGGCGACCAA 1084
QY 251 AACCAATGATATTGCTGTGATGGGACGACTACTGCAACAGTTTGTACACAAGCCATTGT 310
DB 1085 GACCAACGACATCGGGGTGACGGCAACCAACCGCAGCCGCTGCTGCCAGGCGCTGT 1144
QY 311 TCATGAAGGACTAAATAATGTCACAGAGGTGCTTAATCCAATTGTTATCCGTCGAGGAT 370
DB 1145 CCGGAGGCGCTGCGCAACGTTGCGCGCCCGGCGCTTCCCGCGCCCGCTGGAAGAGGGAT 1204
QY 371 TGAACACGACAGCAGCAGCTGTTGAAGCCTTGAAGCCATTGCTCAACCTGTATCTGG 430
DB 1205 CGAGCGCGCGCTCGCGCGCTGTCGCGCGAGCTGCTGACACCGCGCCCGCTGATCGACGA 1264

QY 431 CAAGGAAGCTATTGCTCAGGTGCTGCAGTATCATCATCAGCTCTGAAAAAGTTGGAGATG 490
DB 1265 CAAGTCGACATCGCGCGCTCGCGCGCTCTCCGCGCAGGACAAAGCAGTCTCGGAGCT 1324
QY 491 TATCTCAGAAAGTATGGAGCGTGTGGGCAACGATGGTGTGATACCATCGAAGATCTCG 550
DB 1325 CATCGCGAGCGGTGGCAAGGTGCGCAAGGACGGTGTCTATCAACGCTCGAGAGTCCAA 1384
QY 551 AGGTATGGAACACAGAACTTGAAGTGTTCAGAGCATGCAATTTGACCGTGGTTACTCTGC 610
DB 1385 CACCTTCGGTGTGACCTTGGACTTCACGAGGGCATGGCTTCGACAAAGGGTCACTGTCT 1444
QY 611 TCAATACATGTCACAGCAATGAAAAATGTTGTCAGACCTTGAAGAAACCCATTATCTT 670
DB 1445 CCGGTACATGTTGACCGACGAGCGTATGAGGCGCTCTCGACGACCCGCTACATCT 1504
QY 671 AATCAGGATAAAAAGTGTCAACATCAAGACATTTTGGCAGCTACTTTGAGAGATTTCT 730
DB 1505 GATCCACCGAGGCAAGATCGGTTTCGATCCAGACCTGTCTGCCGCTCTCGAGAGGTCAT 1564
QY 731 TAAAC-----CAACCGTCCATTACTCATTTATTCAGATGATGTGGATGTGAAGCACT 784
DB 1565 CGAGCGGGTGGCTCCAAGCCGCTCTGATCATCCCGAGGACGTCGAGGGCGAGGCCCT 1624
QY 785 TCCAAACCTTGTCTTGAACAAGATTCTGGTACTTTCAATGTGGTGTCTGTCACAAAGCGCC 844
DB 1625 GTCGACCTGTTGGTCAACAAGATCCGCGCACGTTCAACGCGCTCGCGCTCAAGCGGCC 1684
QY 845 AGGATTTGGTATCTGCTAAAGCTATCTGTTGAAGACATTTGCTATCTTGACAGGTGGTAC 904
DB 1685 CGGCTTCGGTTCACCGCGCAGCGATGCTCGCGACATTTGGCACCTCACCGGTGCCAC 1744
QY 905 AGTGATTACAGAGGATCTAGGACTTGAATTAAGACATGCTACAATGACAGCCCTTGGACA 964
DB 1745 CGTATCGCGGAGGAGTGGCGCTCAAGCTCGACGCGGCTCTGGAGCTCTGGGCAC 1804
QY 965 GGCTCTAAGATTACAGTTGATAAGATAGCAGACAGTAATTTGAAGTTTCAAGAGTTTC 1024
DB 1805 CGCGCGCGCGTCAACGTCACCAAGGACGACAGCACCATCGTGGACCTGGAGAGGACGC 1864
QY 1025 AGAAGCTATTGCTAACCGTATTTGCACTGATTAATTCGGAATTTAGAAACAACAATTCGA 1084
DB 1865 CGAGGACCTTCAGGCGCGCTGCGCCAGATCAAGCCCGAGATTCGAGTCGACGACCTCGA 1924
QY 1085 CTTTCACCGTGAAGAACTACAGAAACGTTTGGCGAAATTTAGCTGTGGTGTAGCTGTAT 1144
DB 1925 CTGGGACCGCGAGAGCTCCAGGAGCGCTCGCCAAAGCTGGCCGCGCGCTCTGGCTGAT 1984
QY 1145 CAAAGTAGGAGCTCCACAGACAGACAGCTTTTAAAGAAATGAAACTTCGATTGAGGATGC 1204
DB 1985 CGCGCTCGCGCGCGCACCGAGTCAAGGAGCGCAAGCACCCTCTGGAGAGCGC 2044
QY 1205 TCTAAATGCTACAGTGCAGCGCTTGAAGAGGATTCGTTGCTGTGTGGTGGAAACAGCACT 1264
DB 2045 CATCTCCGCGACCGCGCGCTGCGAGAGGGCATCTCTCCGCTGTGGTGGCTCGCGCT 2104
QY 1265 TAT---TACGGTTATTGAAAGTAGCAGCTTTCAGCTTGGAGCGCGATGCTACTGG 1321
DB 2105 GGTCCACCGCTCAAGTCTTGGACGACAACCTCGGCGCGCAGCGCGAGCGCCACCGG 2164
QY 1322 ACGTAAACATTGCTTCGTGCTTAGAAGAGCGCTGTAGCTCAAAATGCTTTAAATGCTGG 1381
DB 2165 TGTCTCGCTGCTCGCGCGCGCTGCGAGCGCTGCGCTGATCGCGAGAGAGCGCGG 2224
QY 1382 GTACGAAGGCTCCGTAGTATTGACAAAGTTTAAAGAACAGCCCTGACGAGAACAGGATTA 1441
DB 2225 CTTGAGGCTACGCTCATCACCAACCAAGTGGCGGAGCTCGACAAGGCGCGAGGCTTCAA 2284
QY 1442 TGCTGCAACAGGTGAGTGGTGTGATGATTAATAACAGGAATCATTTGACCCCTGTCAAGT 1501
DB 2285 CGCGCGCGCGTACGCGGACCTGTTCAAGCGCGCGCTCATCGACCCCGCTCAAGT 2344

Db 1201 TCCGCGACCCGCGCGGTCGAGAGGGCATCGTCTCCGGTGGTGGCTCCGCGCT 1256

Search completed: February 15, 2001, 23:17:06
Job time: 8581 sec

GenCore version 4.5
Copyright (c) 1993 - 2000, Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: .February 15, 2001, 22:00:13 ; Search time 925.81 Seconds
(without alignments)
12572.166 Million cell updates/sec

Title: US-09-001-737-7
Perfect score: 1661
Sequence: 1 GAATTCGGCTTCATATGGCA.....TCGGCGGATAAGCCGAATTC 1661

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 7991742 seqs, 3503743858 residues 15983484
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*
10: gb_est10.*
11: gb_est11.*
12: gb_est12.*
13: gb_est13.*
14: gb_est14.*
15: gb_est15.*
16: gb_est16.*
17: gb_est17.*
18: gb_est18.*
19: gb_est19.*
20: gb_est20.*
21: gb_est21.*
22: gb_est22.*
23: gb_est23.*
24: gb_est24.*
25: gb_est25.*
26: gb_est26.*
27: gb_est27.*
28: gb_est28.*
29: gb_est29.*
30: gb_est30.*
31: gb_est31.*
32: gb_est32.*
33: gb_est33.*
34: gb_est34.*
35: gb_est35.*
36: gb_est36.*
37: gb_est37.*
38: gb_est38.*
39: gb_est39.*
40: gb_est40.*
41: em_estba.*
42: em_estfun.*
43: em_esthum1.*

44: em_esthum2.*
45: em_esthum3.*
46: em_esthum4.*
47: em_esthum5.*
48: em_esthum6.*
49: em_esthum7.*
50: em_esthum8.*
51: em_esthum9.*
52: em_esthum10.*
53: em_esthum11.*
54: em_esthum12.*
55: em_esthum13.*
56: em_esthum14.*
57: em_esthum15.*
58: em_esthum16.*
59: em_esthum17.*
60: em_esthum18.*
61: em_esthum19.*
62: em_esthum20.*
63: em_estin1.*
64: em_estin2.*
65: em_estin3.*
66: em_estin4.*
67: em_estov1.*
68: em_estov2.*
69: em_estpl1.*
70: em_estpl2.*
71: em_estpl3.*
72: em_estpl4.*
73: em_estpl5.*
74: em_estro1.*
75: em_estro2.*
76: em_estro3.*
77: em_estro4.*
78: em_estro5.*
79: em_estro6.*
80: em_estro7.*
81: em_estro8.*
82: em_estro9.*
83: em_estro10.*
84: em_estro11.*
85: em_estro12.*
86: em_estro13.*
87: gb_est41.*
88: gb_est42.*
89: gb_est43.*
90: gb_est44.*
91: gb_est45.*
92: gb_est46.*
93: gb_est47.*
94: gb_est48.*
95: gb_est49.*
96: gb_est50.*
97: gb_est51.*
98: gb_est52.*
99: gb_est53.*
100: gb_est54.*
101: gb_est55.*
102: gb_est56.*
103: gb_est57.*
104: gb_est67.*
105: gb_est68.*
106: gb_est69.*
107: gb_est70.*
108: gb_est71.*
109: gb_est72.*
110: gb_est73.*
111: gb_est74.*
112: em_esthum21.*
113: em_esthum22.*
114: em_esthum23.*
115: em_estom1.*
116: em_estom2.*

117: em_estpl6:*
 118: em_estpl7:*
 119: em_estpl8:*
 120: em_estpl9:*
 121: em_estpl10:*
 122: em_estpl11:*
 123: em_estpl12:*
 124: em_estpl13:*
 125: em_estpl14:*
 126: em_estpl15:*
 127: em_estpl16:*
 128: em_estpl17:*
 129: em_estpl18:*
 130: em_estpl19:*
 131: em_estpl20:*
 132: em_estpl21:*
 133: em_estpl22:*
 134: em_estpl23:*
 135: em_estpl24:*
 136: em_estpl25:*
 137: em_estpl26:*
 138: em_estpl27:*
 139: em_estpl28:*
 140: em_estpl29:*
 141: em_estpl30:*
 142: em_estpl31:*
 143: em_estpl32:*
 144: em_estpl33:*
 145: em_estpl34:*
 146: em_estpl35:*
 147: em_estpl36:*
 148: em_estpl37:*
 149: em_estpl38:*
 150: em_estpl39:*
 151: em_estpl40:*
 152: em_estpl41:*
 153: em_estpl42:*
 154: em_estpl43:*
 155: em_estpl44:*
 156: em_estpl45:*
 157: em_estpl46:*
 158: em_estpl47:*
 159: em_estpl48:*
 160: em_estpl49:*
 161: em_estpl50:*
 162: em_estpl51:*
 163: em_estpl52:*
 164: em_estpl53:*
 165: em_estpl54:*
 166: em_estpl55:*
 167: em_estpl56:*
 168: em_estpl57:*
 169: em_estpl58:*
 170: em_estpl59:*
 171: em_estpl60:*
 172: em_estpl61:*
 173: em_estpl62:*
 174: em_estpl63:*
 175: em_estpl64:*
 176: em_estpl65:*
 177: em_estpl66:*
 178: em_estpl67:*
 179: em_estpl68:*
 180: em_estpl69:*
 181: em_estpl70:*
 182: em_estpl71:*
 183: em_estpl72:*
 184: em_estpl73:*
 185: em_estpl74:*
 186: em_estpl75:*
 187: em_estpl76:*
 188: em_estpl77:*
 189: em_estpl78:*

190: gb_gss25:*
 191: gb_gss26:*
 192: gb_gss27:*
 193: gb_gss28:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	243	14.6	645	87	AW224051	AW224051 EST300862
2	239.4	14.4	716	140	D46006	D46006 RICS10372A
3	237.6	14.3	637	87	AW223768	AW223768 EST300579
4	236.2	14.2	668	90	AW398404	AW398404 EST298251
5	227.4	13.7	680	89	AW330455	AW330455 PFOVAFGB3
6	226	13.6	1351	108	BE422296	BE422296 HWM022CF
7	225.2	13.6	709	21	A1486676	A1486676 EST244998
8	222.4	13.4	825	110	BE642158	BE642158 Cr12_5_E2
9	220.4	13.3	598	26	A1895994	A1895994 EST265437
10	214.4	12.9	825	97	AW927061	AW927061 HVSMEG000
11	211.2	12.7	695	106	BE318972	BE318972 NF043C12L
12	210.8	12.7	710	104	BE131653	BE131653 L48-1652T
13	209.4	12.6	612	87	AW216817	AW216817 EST295531
14	208.8	12.6	902	105	BE231179	BE231179 HVSMEG001
15	207.4	12.5	739	93	AW622566	AW622566 EST313366
16	201	12.1	677	108	BE435722	BE435722 EST406800
17	199.2	12.0	706	29	AU092993	AU092993 AU092993
18	197.8	11.9	578	108	BE432288	BE432288 EST398817
19	195.8	11.8	651	106	BE318418	BE318418 NF038802L
20	195.6	11.8	552	108	BE434127	BE434127 EST405205
21	195.2	11.8	730	137	BE901181	BE901181 60167579
22	195.2	11.8	819	110	BE643470	BE643470 Cr12_8_O1
23	194	11.7	621	97	AW934661	AW934661 EST353553
24	192.8	11.6	591	108	BE433229	BE433229 EST395758
25	191.8	11.5	557	93	AW618606	AW618606 EST320592
26	190.8	11.5	970	28	AU010446	AU010446 AU010446
27	190	11.4	546	39	AW030594	AW030594 EST273849
28	188.8	11.4	634	97	AW934424	AW934424 EST360267
29	186.8	11.2	734	28	AJ398447	AJ398447 AJ398447
30	184.4	11.1	491	108	BE435362	BE435362 EST406440
31	181.2	10.9	600	108	BE432172	BE432172 EST398701
32	181.2	10.9	886	135	BE799053	BE799053 601588241
33	180.8	10.9	584	38	AW004169	AW004169 701515092
34	180.2	10.8	494	108	BE471174	BE471174 WHE0285_G
35	180	10.8	840	135	BE779875	BE779875 601467908
36	180	10.8	1119	134	BE036873	BE036873 MP07G03_M
37	179.2	10.8	773	109	BE535246	BE535246 601058776
38	178.2	10.7	648	110	BE580002	BE580002 kg35f01.y
39	178	10.7	831	40	AW107008	AW107008 um18d11.y
40	177.2	10.7	609	137	BE918978	BE918978 FMI_2_E09
41	177.2	10.7	700	91	AW506070	AW506070 GEI379_G1
42	176.8	10.6	572	109	BE496298	BE496298 NXCI_022
43	176.4	10.6	578	108	BE432591	BE432591 EST399120
44	176	10.6	528	37	AV525715	AV525715 AV525715
45	176	10.6	828	106	BE300379	BE300379 600944444

ALIGNMENTS

RESULT 1
 AW224051
 LOCUS EST300862 tomato fruit red ripe, TAMU Lycopersicon esculentum CDNA
 DEFINITION clone cLE14H9, mRNA sequence.
 ACCESSION AW224051
 VERSION AW224051.1
 KEYWORDS EST.
 SOURCE tomato.
 645 bp mRNA
 EST 07-DEC-1999

ORGANISM	Lycopersicon esculentum Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon. 1 (bases 1 to 645)
REFERENCE	
AUTHORS	Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F., Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Ronning,C.M. , Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J. Generation of ESTs from tomato fruit tissue Unpublished (1999)
TITLE	
JOURNAL	
COMMENT	Contact: David Frisch Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 4366 Fax: 864 656 4293 Email: dfrisch@CLEMSON.EDU 5 prime sequence.
FEATURES	Location/Qualifiers 1..645 /organism="Lycopersicon esculentum" /cultivar="TA496" /db_xref="taxon:4081" /clone="cLEM14H9" /clone_lib="tomato fruit red ripe, TAMU" /tissue_type="pericarp" /dev_stage="red ripe (7-20 days post-breaker)" /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; supplier: Giovannoni; Fruit were tagged at the breaker stage (first sign of lycopene accumulation on the blossom end of the fruit) and harvested 7 days post-breaker (fully red-ripe), 10 days post breaker, and 20 days post-breaker (over-ripe). 20 day fruit which showed external or internal signs of pathogenesis were discarded. Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."
BASE COUNT	205 a 166 g 172 t

	Query Match	14.6%	Score 243:	DB 87:	Length 645:	
	Best Local Similarity	61.1%	Pred. No. 1.8e-57:			
	Matches 393:	Conservative	0:	Mismatches 250:	Indels 0:	Gaps 0:
QY	607	TGCTCTAAATCATGTGCACAGACAATGAAATAATGGTTGCGAGACCTTGAAATCCCAATTA	666			
DB	2	TCCTCTCTTACTTGTGTACCGACAGTGCAGAAATATGCTCCGTTGAATATGAGNACTGTGAAGT	61			
QY	667	TCCTTAATCAGCGGATAAAAAAGTGTCAACATCCAAAGACATTTGGCCACTACTTTCAGGAAG	726			
DB	62	TGCTACTGGTTGTATAAAAAGATACAAATGCAAGAGATCTTGTAAATGCTGCTGGAAGATG	121			
QY	727	TTCTTAAACCAACCGTCCACTTACTCATATTATGCAGATGATGGATGGTGAAGCACTTC	786			
DB	122	CTATCAGAAATGGTTACCCAAATTTTAATTAATGCTGAAGATATTTGACGAGAGCTTTGG	181			
QY	787	CAACCCCTTGCTTTGAACAAGATTGCTGGTACTTTCAATGGTGTCTGTCAAAGCGCCAG	846			
DB	182	CAACTCTTGTGTCAATAAGCTTAGAGGTGCTTGAAGGTGCGTGCACCTTAAAGCTCCTG	241			
QY	847	GANTTGGTGATCGTCGTAAGCCTATGCTTGAAGACATTCCTATCTTCACAGGTGGTACAG	906			
DB	242	GTTTTGGTGAGCAAAAAGCCAGTATCTTGATGACATAGCAACCCCTTACTGGAGGCCACTG	301			
QY	907	TGATTACAGAGGATCTAGGACTTTGAATTAAGAATGCTACAATGACAGCCCTTGGACAGG	966			
DB	302	TTATTAGGGAGGAGCTTGGCCCTTACCTTTGGACAAGGCTGACAAGGAAGTCTAGTCTATG	361			
QY	967	CTGCTAAGATTACAGTTGATTAAGAATAGACAGATAATTGTTGAAGGTTTCAGGAAGTTTCAG	1036			
DB	362	CTGCTAAGTAGTGTCTGACTAAGGATGCCACTACAATTTGGTGTGGTGAAGCACTCAGG	421			
QY	1027	AAGCTATTGCTAACCGTATTTCGCACTGATTAATTCGCAATTTAGAAACAACAACTTCTCACT	1086			

Db	422	AAGCAGTCAACAACACGTTGTCACAGATTAAAAACCTGATAGAGGTCACAGATCAAGATT	481
QY	1087	TTGACCGTGAAAAACTACAAGACGTTTGGCGAAATTAGCTGGTGGTAGCTGTTATCA	1146
Db	482	ATGAAAGGAAAGCTTAATGAAGAAATTCCTAAGTTATCAGGAGGTGGCTGTCTATC	541
QY	1147	AAGTAGGAGCTCCAAACAGACACAGCTTTAAAGAAATGAACTTCGCAATGAGGATGCTC	1206
Db	542	AGGTGGAGCTCAACCTGAACCTGAATGAAGGAGAGAACTTAGAGTAGAAGATGCTC	601
QY	1207	TAAATGCTACAGCTGCAGCGTTGAAGAAGTATCGTTGCTGG	1249
Db	602	TCAATGCAACAAGGCGCTGTTGAGGAAGTATTTGTTTGG	644

RESULT	2
D46006	
LOCUS	D46006 716 bp mRNA EST 09-MAR-2000
DEFINITION	RICS10372A Rice green shoot Oryza sativa cDNA, mRNA sequence.
ACCESSION	D46006
VERSION	D46006.2 GI:71212768
KEYWORDS	EST.
SOURCE	Oryza sativa.
ORGANISM	Oryza sativa
REFERENCE	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.
AUTHORS	1 (bases 1 to 716)
TITLE	Sasaki, T., Miyao, A. and Yamamoto, K.
JOURNAL	Rice cDNA from callus 1995
COMMENT	Unpublished (1995) On Mar 9, 1995 this sequence version replaced gi:699715. Contact: Takuji Sasaki National Institute of Agrobiological Resources Rice Genome Research Program 2-1-2 Kannondai, Tsukuba Ibaraki, Japan 305 Tel: 0298-38-7441 Fax: 0298-38-7468 Email: tsasaki@abr.affrc.go.jp PROJECT = 'RGP' Sequence updated (01-Mar-2000). Location/Qualifiers 1..716 /organism="Oryza sativa" /strain="Nipponbare" /db_xref="taxon:4530" /clone_lib="Rice green shoot" /note="Green shoot (8 days old)" 222 a 120 c 198 g 175 t 1 others

FEATURES	source	
BASE COUNT	222 a 120 c 198 g 175 t 1 others	
ORIGIN		
Query Match	14.4%; Score 239.4; DB 140; Length 716;	
Best Local Similarity	58.6%; Pred. No. 1.9e-56;	
Matches 414; Conservative	0; Mismatches 292; Indels 0; Gaps 0;	
QY	550 GAGGTATGGAACAGAACTTGAAGTGGTTGAAGGCATGCAATTTGACCGTGGTTACCTGT	609
Db	5 GAAGTCTGAGATAATCTCTATGTTGTCGAGGGAATGTCAGTTTGAACGTGGTTATATCT	64
QY	610 CTCAATACATGGTTCACAGACAANTGAANAATGTTGCAGACCTTGAAACCCCATTTATCT	669
Db	65 CCCCGTATTTTGAATCTGACGACGAGAGAAATGTCGAGAAATATGAGAACTGCAANCTTC	124
QY	670 TAATCAGCGATAAAAAAGTGTCAAACTCCAAAGACATTTTGGCCACTACTTTGAGGAAGCTTC	729
Db	125 TTTTGGTGGAACAAAATCACCACCCGAGGATCTTATCAATGTTTGGAGGAGGCCA	184
QY	730 TTAACAAACACCGTCCATTACTTATTCAGATGATGGTGGTGAAGACACTTCCAA	789
Db	185 TAAGAGTGTGATCCCAATCTTCATATTCCTCAGGATATTTGACGACGAGCGCTTGGCTA	244

Db	484	GCACATCAAAATTTGGTGTAGTGACACTCAGGAAGCAGTCAACAACAGGTGGTCACAG	543
QY	1053	ATTAATAATCGCAATATTAGAAACAACTCTGACTTTGACCGGTGAAATAACTACAAGACGT	1112
Db	544	ATTAAAACTGTATAGAGGCTCGAGATCAAGATTATGAAAGGAAAGCTTAATGAAGA	603
QY	1113	TTGGCGAAATTAAGTCTGGTGTAGCTGTATCAAGATAGGAGCTCCAAACAGACAGCT	1172
Db	604	ATTGCTAAGTTATCAGGAGGTGGCTGTCATACAGCTTGGAGCTCAAACTGAAACTGAA	663
QY	1173	TTAAA	1177
Db	664	TTGAA	668

RESULT	5
AW330455	
LOCUS	AW330455 680 bp mRNA EST 31-JAN-2000
DEFINITION	PFOVAFCB331SK Onchocerca volvulus adult female cDNA clone following ivermectin (SAW98PF-OVAF) Onchocerca volvulus cDNA clone PFOVAFCB331 5', mRNA sequence.
ACCESSION	AW330455
VERSION	AW330455.1 GI:6826808
KEYWORDS	EST.
SOURCE	Onchocerca volvulus.
ORGANISM	Onchocerca volvulus
REFERENCE	Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea
AUTHORS	Onchocercidae; Onchocerca.
TITLE	1 (bases 1 to 680)
JOURNAL	Fischer,P. and Williams,S.A.
COMMENT	Genes Expressed in adult female stage of Onchocerca volvulus following treatment with ivermectin Unpublished (1999)
FEATURES	<p>Location/Qualifiers</p> <p>1..680</p> <p>/organism="Onchocerca volvulus"</p> <p>/db_xref="taxon:6282"</p> <p>/clone="PFOVAFCB331"</p> <p>/clone_lib="Onchocerca volvulus adult female cDNA following ivermectin (SAW98PF-OVAF)"</p> <p>/sex="female"</p> <p>/dev_stage="adult"</p> <p>/lab_host="XLI-Blue MRF"</p> <p>/note="Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2: Xho I; Filarial nematode parasite of humans. Two adult female worms of Onchocerca volvulus were isolated from one consenting patient from western Uganda. The patient was treated 28 hours and 7 month before nodulectomy with single dose of 150 mg/kg ivermectin. Adult female worms were quick frozen. Their mRNA was converted to double-stranded cDNA using reverse transcriptase and oligo(dT) followed by RNase H and DNA pol I. The library has 2.9 x 10E5 independent recombinants and the average insert size is ~800bp. The library was constructed by Peter Fischer with worms provided by Dr. Dietrich W. Buttner. The library is available from Dr.Steven A. Williams (U.S.A.) or Dr. Peter Fischer (Germany), email: genome@smith.edu, Fischpe@aol.com"</p>
BASE COUNT	230 a 93 c 151 g 204 t 2 others
ORIGIN	

Query Match	13.7%;	Score	227.4;	DB	89;	Length	680;
Best Local Similarity	61.2%;	Pred. No.	4.6e-53;				
Matches	385;	Conservative	0;	Mismatches	241;	Indels	3;
							Gaps

QY 611 TCAATACATGGTCACAGACAATGAAATAATGTTGCGAGACCTTTGAAACCCCAATTTATCTT 670
 Db 1 TCGTATTTTATACAAATAATGAGAAATGATTTGGAGCTTGATGATCCATATCTCT 60
 QY 671 AATCAGGATATAAAGAGTGTCAACATCAAGACATTTTCCACACTTCTTGAGGAAGTTCT 730
 Db 61 AATACAGAGAAATAAATCTTAATATTATTCAGGCTTTTCTTCTTATTTCTTGAAGCTTGT 120
 QY 731 TAAACCAACCGTCCATTTACTTATTCATATTTCAGATGATGTGGATGTGAAGCACTTCCCAAC 790
 Db 121 TAAGTCTGTAACCTTTACTTATTCATTCGCAAGACATTTGAGGTTGAAGCATTAAGTAC 180
 QY 791 CTTGCTTTGAACAGATTCGTTGATCTTTCATTTGCTGTCGCAAGCCGAGGATT 850
 Db 181 TTTAGTTATTAAAGTTCGTTGAGGCTCTAAAGTTTGTGCAAGTGAAGCTCCAGGTTT 240
 QY 851 TGGTATCTGCTGAAGCTATGCTTTGAAGACATTTGCTATCTTTCACAGGTTGTA---CAGT 907
 Db 241 TGGTGATAGAAGAAAGGAGATGCTTTGAAGATATAGCAGCTTTTAACTTAATGCTAAAGTATCT 300
 QY 908 GATTACAGAGGATCTAGGACTTTGAATTTAAAGATGCTACAATGACAGCCCTTGGACAGGC 967
 Db 301 CATAAAGATGACTTGGATTTAAATGCAAGACCTTAACCTTTGAAGACCTTTGGCATTCG 360
 QY 968 TGTAAAGATTACAGTTGATAAAGATAGACAGTAAATTTGTTGAAGTTTCAGGAAGTTTCA 1027
 Db 361 TAAATATGTTAAATACACTAAAGATTAATTAATTAATTTGTTAGCAAAATAGAGTTACTGA 420
 QY 1028 AGCTATTGCTACCGTATTGCTACTGATTAATTCGCAATTTAGAACACAACTTTCGACTT 1087
 Db 421 CAGAGTAAAGCTAGATTTGACAGATTAATTAATTTCAATTTGAGTCTTCAACTTTCTGATTA 480
 QY 1088 TGACCGTGAATAACTACAAAGACGTTTGGCAAAATAGCTGCTGTTGTTAGCTTTATCAA 1147
 Db 481 TGTAAAGAGAAATTAAGAGCGTTTACCAAAATATACAGTTGTTGCTTGTACTATAA 540
 QY 1148 AGTAGAGCTTCCAAGACAGACGCTTTTAAAGAAATGAAACTTCCGATTTGAGGATGCTCT 1207
 Db 541 AGTTGTTGAGCACCTGAAATTAAGATTTAAAGAACCTTAAGATAGGTCGAGGATGCTCT 600
 QY 1208 AATGCTACAGCTGACCGCTTTGAAGAAG 1236
 Db 601 GCACGCCACAAGAGCTGCAATTTGAGGAAG 629

RESULT 6
 BE422296 LOCUS BE422296 1351 bp mRNA EST 24-JUL-2000
 DEFINITION HW022CF.04r ITEC HWM Barley Leaf Library Hordeum vulgare cDNA
 clone HW022CF.04, mRNA sequence.
 ACCESSION BE422296
 VERSION BE422296.1 GI:9420139
 KEYWORDS EST.
 SOURCE barley.
 ORGANISM Hordeum vulgare
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliophyta; Liliopsida; Poales; Poaceae; Hordeum.
 REFERENCE 1 (bases 1 to 1351)
 AUTHORS Anderson, O.A., Appels, R., Bailey, P., Blake, T., Close, T., Cloutier,
 S., Dubcovsky, J., Feuillet, C., Gale, M., Graner, A., Gustafson, P.,
 Herrmann, R.G., Holton, T., Jacquemin, J.M., Jia, J., Joudrier, P.,
 Langridge, P., Lazo, G.R., Lin, J.J., McGuire, P., Ogihara, Y.,
 Pecchioni, N., Qaibet, C., Schuch, W., Selvaraj, G., Shariflou, M.,
 Sorrells, M., Warburton, M. and Wenzel, G.
 International Triticeae EST Cooperative (ITEC): Production of
 Expressed Sequence Tags for Species of the Triticeae
 Unpublished (2000)
 CONTACT: Wenzel G
 TU Muenchen, Lehrstuhl fuer Pflanzenbau und Pflanzenzuchtung
 Am Hochanger 2, D-85350 Freising-Weihenstephan GERMANY
 Fax: 49 08161 71 5173
 Email: wenzel@mm.pbz.agrar.tu-muenchen.de

International Triticeae EST Cooperative (ITEC)
 http://wheat.pw.usda.gov/genome.
 Location/Qualifiers
 1. 1351
 /organism="Hordeum vulgare"
 /cultivar="Barke"
 /db_xref="taxon:4513"
 /clone="HW022CF.04"
 /clone_lib="ITEC HWM Barley Leaf Library"
 /tissue_type="leaf"
 /dev_stage="14 day old"
 /note="Vector: pBluescriptSK(-); 850 bp average insert
 size."
 BASE COUNT 376 a 281 c 372 g 317 t 5 others
 ORIGIN

Query Match 13.6%; Score 226; DB 108; Length 1351;
 Best Local Similarity 58.6%; Pred. No. 1.5e-52;
 Matches 391; Conservative 0; Mismatches 276; Indels 0; Gaps 0;

QY 632 TGAATAAATGGTTGAGACCTTTGAAACCCCAATTTATCTTAATCACGGATATAAAGTGTG 691
 Db 119 TGAGAAATGACCCAGGTACGAGAACTGCAAGCTGCGCTTGGTTGACAGAAATACAC 178
 QY 692 AAACATCCAAAGATTTTGGCCACTTCTTGAGGAAGTTCTTAAACCAACCCGCTCCATCT 751
 Db 179 CAACGACGGATCTTATCAATGTTCTGGAGGAGCCATTTAGGGTCAATACCACTCT 238
 QY 752 CATATTGTCAGATGATGTTGGATGGTGAAGCACTTCCAAACCTTGTCTTGAACAAGATTG 811
 Db 239 GATCATTTGCTGAGGATTTTGAAGAGGCTCTTGAACCCCTTGTGTCAACAAGCTAAG 298
 QY 812 TGSTACTTTCAATGTTGTTGCTGTCAAACGCCAGGATTTGGTGATCGCTGAAGCTAT 871
 Db 299 AGTCTTTTGAATAATCTGTGCTATCAACGCCCTTGGTTTGGAGAGCGCAACCCAGTA 358
 QY 872 GCTTGAAGACATTTGCTTCTTACAGAGTGTGTACAGTGTATACAGAGATCTAGGACTTGA 931
 Db 359 CTTGGAGACATTTGCCATCTTCAACCGGAGAACTGTTATCAGAGAGAGGTTGGACTCAC 418
 QY 932 ATTAAGAATGCTACATGACAGACCCCTTGGAGAGGCTGCTAAGATTACAGTTGATAAGA 991
 Db 419 ACTTGACAGGAGATTAACACAGTTCTAGGAACGGCTGCAAGAGTTGTCTTACAAAGA 478
 QY 992 TAGCACAGTAATTTGTTGAAGTTTCAGGAAGTTTCAAGAGCTATTGCTTAACCGTATTC 1051
 Db 479 GTCGACACAATAGTTGTTGTTGAGCAGCACCCAGGAAGTGAAGTGAAGAGGTTGCACA 538
 QY 1052 GATTAATCGCAATTTAGAAACCAACTTCTGACCTTTTGACCTTGAAACCACTACAAGAAG 1111
 Db 539 GATCAAAATCTCATTTAGGTAGCAGAGCAACACTACGAGAGGAGAAACTCAATGAGAG 598
 QY 1112 TTTGGCGAAATTTAGCTGTTGGTGTAGCTGTTATCAAGTAGGAGCTTCCAAACAGACAGC 1171
 Db 599 GATTGCAAGAGCTGCGCGGTTGTTGCTGTTATTGAGTTGGGAGGACCAACAAGAACTGA 658
 QY 1172 TTTAAAGAAATGAACCTTCGCATTTGAGATGCTCTTAATGCTACAGCTGACGCGCTTGA 1231
 Db 659 ACTTGAAGAGAGAGTGTGCGAGTTGAGATGCTCTTANACGCAACCAAGCTTGCCTTGA 718
 QY 1232 AGAAGTATCGTTGCTGTTGGTGGTGAACACACTATTATACGTTATTGAAAAAGTAGCAGC 1291
 Db 719 GGAAGTATGTTGTTGGTGGGAGGCTGACACTCTTTTGGAGCTGGCTGCTAAGTTGATGC 778
 QY 1292 TCTTGAAG 1298
 Db 779 CATCAAG 785

RESULT 7
 A1486676 LOCUS A1486676 709 bp mRNA EST 29-JUN-1999

```

DEFINITION EST244998 tomato ovary, TAMU Lycopersicon esculentum cDNA clone
            cLED11F20, mRNA sequence.
ACCESSION  A1486676
VERSION     A1486676.1 GI:4382047
KEYWORDS    EST.
SOURCE      tomato.
ORGANISM    Lycopersicon esculentum
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids
            I; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1 (bases 1 to 709)
AUTHORS     Alcalá J., Vrebalov J., White, R., Matern, A.L., Vision, T., Holt, I. E.,
            Liang, F., Upton, J., Ronning, C.M., Craven, M.B., Fujii, C.Y., Bowman
            , S.D., Nierman, W., Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley
            , C.L., and Giovannoni, J.
            Generation of ESTs from tomato carpal tissue
            Unpublished (1999)
            Contact: David Frisch
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 4366
            Fax: 864 656 4293
            Email: dfrisch@CLEMSON.EDU,
            Location/Qualifiers
FEATURES             source
                     1..709
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLED11F20"
                     /clone_lib="tomato ovary, TAMU"
                     /tissue_type="carpel"
                     /dev_stage="5 days pre-anthesis to 5 days post-anthesis"
                     /lab_host="XL1-Blue MRF"
                     /note="vector: pBlueScript SK(-); Site.1: EcoRI; Site.2:
                     XhoI; cLED - Tomato Carpel EST Library. OligodT-primed and
                     directionally cloned cDNA in vector Lambda ZAP II with 5'
                     and 3' ends located at the EcoRI and XhoI sites,
                     respectively."
BASE COUNT      219 a 110 c 193 g 187 t
ORIGIN
13.6%; Score 225.2; DB 21; Length 709;
Best Local Similarity 58.8%; Pred. No. 1.9e-52;
Matches 389; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

QY 478 AAGTTGGAGACTATATCTCAGAACGTATGGAGCGTGGCGCAACGATGCTGTGATTACCA 537
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 48 AAGTAGGAGTATGATGCTGAAGCCATGAGCAAGGTTGGGAAGGATGTTGTGACAC 107

QY 538 TCGAAGAAATCTCGAGGTATGGAACACAGAACTTGAAGTGTGTTGAAGGCATGCAATTTGACC 597
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 108 TAGAGGAGGTAAAGTCTGAAACAGCTGCGGTGTTGAGGAGTGAATTTGACC 167

QY 598 GTGGTTACCTGCTCAATACATGATGTCAGACAAATGAAAAAATGTTGTCAGACCTTGAAA 657
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 168 GTGGTTATGCTCTCTCTTACTTGTGTCAGACAGTGAAGAAATGTCCTTGAATATGAGA 227

QY 658 ACCCATTTATCTTAATCAGGATAAAGTGTCAACATCCAGACATTTTGCCATAC 717
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 228 ACTGTAAGTTGCTACTGTTGATAAAAAAGATAACAAATGCAAGAGATCTTCTTAATGTCC 287

QY 718 TTGAGGAAGTTCTTAAACCAACCGTCCATCTACTTACTTATTTGAGATGATGTTGATGTTG 777
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 288 TGGAGATGCTATCAGAAATGGTTACCCCAATTTTAATATTCTGTAAGATATTGAGCAGG 347

QY 778 AAGCACTTCCAAACCTTGTCTTTGAACAGATTCGTGCTACTTCAATGTGTTCTCTGPCA 837
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 348 AAGCTTTGGCAACTCTTGTGTTGTAATAGCTTAGAGGTGCTTGAAGGTCGCTGCACATTA 407

QY 838 AAGCCCAAGGATTTGGTGATCGTCTGTAAGCTATGCTTTGAAGACATTTGCTATCTTGACAG 897
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

DB 408 AAGTCTCTGGTTTGGGTGAGCGAAAAAACCCAGTATCTTTGATGACATACCAACCCTTACTG 467
QY 898 GTGGTACAGTGAATACAGAGGATCTAGGACTTTGAATTTAAAGATGCTACAAATGACAGCCC 957
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 468 GAGGCACTGTTATTAGGAGGAGGCTTGGCCCTTACCTTGGACAAGGCTGACAAGGAGTTTC 527

QY 958 TTGGACAGGCTGCTAAGATTACAGTTGATAAAGATAGCACAGTAAATTTTGAAGGTTTCAG 1017
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 528 TAGGTCATGCTGCTAAAGTAGTGTGCTAAGGATGCCACTACAAATTTGTTGTTGATGTTA 587

QY 1018 GAAGTTCAAGAGCTATTGCTTAACCGTATTGCACCTGATTAAATTCGCAATTAGAACAAACAA 1077
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 588 CCACTCAGGAAGCAGTCAACAAACGTTGTCACAGATTAAACCTGATGAGCGCTGCAG 647

QY 1078 CTTCTGACTTTGACCGTGAAGAAACATACAAAGACGTTTGGCGAAATTTAGCTGTGTGTAG 1137
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 648 ATCAAGATTATGAAAGGAAAGCTAAATGAAGAATTGCTTAATTTATCAGGAAGTGTGG 707

QY 1138 CT 1139
DB 708 CT 709

RESULT      8
BE642158    BE642158    825 bp    mRNA    EST    01-SEP-2000
LOCUS       CRI2_5_E21_SP6 Ceratopteris Spore Library Ceratopteris richardii
DEFINITION cDNA clone CRI2_5_E21 5', mRNA sequence.
ACCESSION   BE642158
VERSION     BE642158.1 GI:9959831
KEYWORDS    EST.
SOURCE      Ceratopteris richardii.
            Ceratopteris richardii
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Filicophyta;
            Filicopsida; Filicales; Pteridaceae; Ceratopteris.
REFERENCE   1 (bases 1 to 825)
AUTHORS     Chatterjee, A., San Miguel, P., Stout, S.C., Banks, J., and Roux, S.J.
TITLE       Expressed sequence tags of cDNA clones from a C. richardii library
            Unpublished (2000)
COMMENT     Contact: Roux SJ
            Section of Molecular Cell and Developmental Biology
            University of Texas
            Biology Building, Room 16, Austin, TX 78712, USA
            Tel: 512 471 4238
            Fax: 512 232 3402
            Email: sroux@uts.cc.utexas.edu
            Plate: CRI2.5 row: E column: 21
            Seq primer: SP6.
            Location/Qualifiers
            1..825
            /organism="Ceratopteris richardii"
            /cultivar="Brogn"
            /db_xref="taxon:49495"
            /clone="CRI2_5_E21"
            /clone_lib="Ceratopteris Spore Library"
            /tissue_type="Gametophyte"
            /cell_type="Spore"
            /dev_stage="20 hours after germination initiation"
            /note="vector: pCMVSPORT6; EST sequence from cDNA library.
            cDNA library constructed from mRNA isolated from C.
            richardii spores that had developed for 20 hours after
            their germination had been initiated by white light."
BASE COUNT      244 a 152 c 229 g 200 t
ORIGIN
13.4%; Score 222.4; DB 110; Length 825;
Best Local Similarity 58.7%; Pred. No. 1.3e-51;
Matches 404; Conservative 0; Mismatches 281; Indels 3; Gaps 1;

QY 1 GAATTCGCTTCATATGCAAAAGAAATCAAAATTTTCAGCAGATCGCGTCTGCCATGG 60
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 55 GAGTCCGGAACATATGCAAGCAAGACATCAGGTTTGGTATTGAGGCCGCGGACCACATGC 114

```

FEATURES		Location/Qualifiers	
source		1. 598	
		/organism="Lycopersicon esculentum"	
		/cultivar="TA496"	
		/db_xref="taxon:4081"	
		/clone="CLEC13D21"	
		/clone_lib="tomato callus, TAMU"	
		/tissue_type="callus"	
		/dev_stage="25-40 days old"	
		/lab_host="XL1-Blue MRF"	
		/note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2: XhoI; supplier: Giovannoni laboratory; cLEC - Cotyledons of seedlings 7-10 days post-germination were excised, cut at both ends and placed on MS medium with no selection. Mixed callus was harvested at 25 and 40 days and included undifferentiated masses. Tomato Callus EST Library"	
BASE COUNT		188 a 97 c 150 g 163 t	
ORIGIN			
Query Match		13.3%; Score 220.4; DB 26; Length 598;	
Best Local Similarity		60.5%; Pred. No. 4.1e-51;	
Matches 362; Conservative		0; Mismatches 236; Indels 0; Gaps 0;	
QY	580	AAGGCGCAATTTGACCGTGTACCTGCTCAATACATGCTCAGACACAATGAAAAA	639
Db	1	AAGGCGCAATTTGACCGTGTACCTGCTCAATACATGCTCAGACACAATGAAAAA	60
QY	640	TGGTTGCAACACCTTGAACACCCATTTATCTTAATCAGCGATAAAAAGTGTCAACATCC	699
Db	61	TGTCCGTTGAATATGAGAACTGTAACTTCTGTTGATATAAAGATAACAAATGCAA	120
QY	700	AAGACATTTTGGCACTTACCTTGAAGAGTTCTTAAACCAACCGTCCATTTACTATTATG	759
Db	121	GAGATCTGTTTAACTGCTCGGAAGATGCTATCAGAAATGTTACCCATTTTAAATATTG	180
QY	760	CAGATGATCGGATGGTGAAGCACTTCCAAACCTTGTCTTGAACAAAGATTCTGGTACTT	819
Db	181	CTGAAGATATTGAGCAGGAAGCTTTGGCAACTCTTGTCTCAATAAGCTTAGAGGTGCCT	240
QY	820	TCAATGCTGTTCTGTCAAAGCCGAGGATTGGTGATGCTGCTGAAGCTATGCTTGAAG	879
Db	241	TGAAGTCTGCTGCACTTAAAGCTCTGTTGGTGAGCGGAAAGCCAGTATCTTTGATG	300
QY	880	ACATTTGCTATCTTACAGGTGTTACAGTATTACAGAGGATCTAGGACTTGAATTAAGAAG	939
Db	301	ACATAGCAACCCCTTACTGGAGGCACTGTTATTAGGAGGAGGCTTGGCTTACTTGGACA	360
QY	940	ATGCTACAATACAGCCCTTGGACAGGCTGCTGAAGATTACAGTTGATAAAGATAGCACAG	999
Db	361	AGGCTCACAAGGAAGTTCTAGGTCATGCTGCTAAAGTAGTGTGCTAGGATGCCACTA	420
QY	1000	TAAATTTGCAAGGTTTCAGGAAGTTTCAGAGGCTATTCCTAACCGTATTGCACTGATTAAT	1059
Db	421	CAATTTGTTGGTATGTTAGTACTAGCAAGAGGCTCAACAAACGCTGTTGCACAGATTAAA	480
QY	1060	CGCAATTTAGAAAACAACCTTCTGACTTTGACCGCTGAAAAACTTACAAGAACCTTTGGCGA	1119
Db	481	ACCTGATAGAGCTCGCAGATCAAGATTATGAAGAAGAAAGCTAAATGAAGAATGCTA	540
QY	1120	AATTAGTGGTGTAGCTGTTATCAAGTAGGAGCTCCACAGACAGACAGCTTTAA	1177
Db	541	AGTTATCAGGAGGTGTGGCTGTTCATACAGGTTGGAGCTCAAACTGAACTGAATGAA	598
RESULT 10			
AW927061		825 bp mRNA EST	
LOCUS		HVSMEG0009G12 Hordeum vulgare pre-anthesis spike EST library	
DEFINITION		HVSMEG0009G12 (white to yellow anther) Hordeum vulgare cDNA clone	
ACCESSION		AW927061	
VERSION		AW927061.1	
		GI:8098803	

QY	61	TGCGCGAGTGTATGTTAGCAGATACCGTCAAGTAAGCTTGGTCTCTAAAGGGCGCA	120
Db	115	TTCAAGAGGTGGAGCAGCTTGCAGATCGGCTCAAGATTACTATGGTCTCAAGGGCGCA	174
QY	121	ATGCTGCTTGAAGAGCTTTGGTCTCCCTTAATTAATGATGACGGGTAAACATTG	180
Db	175	ATGTTGTTATGACAAAGCTTTGGATCCCAAGGTCACCAAGATGGTGTCACTGTAG	234
QY	181	CTAAGAGATCGAATTAGAAGATCATTTTGAACACATGGGAGCAAAATGGTCTCTGAAG	240
Db	235	CTAAGAGCTTGAATTCAGACAGAGTGTGAGACATTTGGAGCAAGTCTGTCAACCAAG	294
QY	241	TGGCTCTTAAACCAATATATTGCTGGTATGGGAGCAGTACTGCAACAGTTTTCACAC	300
Db	295	TAGCGAATGCTACAAATATGACAGAGGGATGCAACTACCTGTGCCACTGTCTTGACAA	354
QY	301	AAGCCATTTGTCATGAGGAGCTAAAAAATGTGACACAGGTGCTAATCAATTTGGTATCC	360
Db	355	GAGCTATATATCTAGGAGTGCANAATCCGTTGCTCGGGGAATGAATGCCATGATTTAC	414
QY	361	GTCGAGCATTTGAACAGACAGCAACAGCTGTTGAAGCCTTGAAGCATTGCTCAAC	420
Db	415	GTCGAGTATTACATGCGCAGTTGAGCTGTTGTGGAACACTTGAAGAGGAGGCGCAAA	474
QY	421	CTGTATCTGGCAAGAGATTTGCTCAGTGTGCTGCAATATCATCACGCTCTGAAAAAG	480
Db	475	TGATAAGTATACATCAGAAGAGATCGCTCAGTTGGTACTATATCTGCTAACGGTGATAGG	534
QY	481	---TTGAGAGTATCTCAGAGCTATGAGCGTGTGGCAGAGTGTGTGATTAACA	537
Db	535	AGATTTGCTGAGCTGATTGTTAAGCTTATGAGAGAGTTGGTAAAGATGGTGTGATACTG	594
QY	538	TGGAAGATCTCGAGTATGGAACAGAACTTGAAGTGTGTTGAAGCATGCAATTTGACC	597
Db	595	TATCTGATGGGAACAATGACCAATGAATGAAGTTGAGGGTATGATGCTGGATA	654
QY	598	GTGGTTACCTGCTCAATACATGCTGACAGCAATGAAAAATGTTGAGAGCTTTGAAA	657
Db	655	CGGCTGCACTGCTCCCTACTGTCATCGATGCTTAAACTCCGAACCTGTGACATGAGAG	714
QY	658	ACCATTTATCTTAATCAGGATAAAAA	685
Db	715	AACCTTGGCAATCTTATTTCGACTACAA	742
RESULT 9			
AW95994		598 bp mRNA EST	
LOCUS		EST265437 tomato callus, TAMU Lycopersicon esculentum cDNA clone	
DEFINITION		CLEC13D21, mRNA sequence.	
ACCESSION		AW95994	
VERSION		AW95994.1	
KEYWORDS		EST.	
SOURCE		tomato.	
ORGANISM		Lycopersicon esculentum	
		Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.	
REFERENCE			
AUTHORS		Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E., Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ahn,S., Ronning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.	
TITLE		Generation of ESTs from tomato callus tissue	
JOURNAL		Unpublished (1999)	
COMMENT		Contact: David Frisch Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 4366 Fax: 864 656 4293 Email: dfrisch@CLEMSON.EDU 5 prime sequence.	

KEYWORDS EST.
SOURCE Barley.
ORGANISM Hordeum vulgare
REFERENCE Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Mangoliophyta; Liliopsida; Poales; Poaceae; Hordeum.
AUTHORS 1 (bases 1 to 825)
Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D., Yu
,T., Anderson,H., Dale,J., Henry,D., Kernodle,S., Palmer,M., Rambo
,T., Saski,C., Schwartzbeck,J., Simmons,J., Choi,D.W., Main,D. and
Wood,T.
TITLE Development of a genetically and physically anchored EST resource
for barley genomics
JOURNAL Unpublished (2000)
COMMENT Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
High quality sequence stop: 825.
FEATURES
source
1..825
/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"
/clone="HVSMEg0009G12"
/clone_lib="Hordeum vulgare pre-anthesis spike EST library
HVCdNA0008 (White to yellow anther)"
/tissue_type="pre-anthesis spike"
/lab_host="SOLR"
/note="vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 242 a 166 c 237 g 179 t 1 others
ORIGIN
Query Match 12.9%; Score 214.4; DB 97; Length 825;
Best Local Similarity 60.4%; Pred. No. 2.2e-49;
Matches 353; Conservative 0; Mismatches 231; Indels 0; Gaps 0;
QY 715 TACTTTGAGGAGTCTTAAACCACCGTCCATTACTTACTTATTCGATGATGCTGATG 774
DB 11 TTCTGGAGGAGCCATAGGGGTCAATACCCAACTCTGATCATTCGTGAGGATATGAGC 70
QY 775 GTGAAGCAGCTTCCAAACCTTCTGTGAACAAGATTCGTGACTTTCATGTTGCTGTG 834
DB 71 AGAGGCTCTTGCAACCTTCTGTCAACAAGCTAAGAGGTTCTTTGAAATCTGTGCTA 130
QY 835 TCAAGGCCAGGATTTGGTGATCGTCGTAAGCTATGCTTGAAGACATTCCTATCTTGA 894
DB 131 TCAAGGCCCTTGGTTTGGAGAGCGCAAGACCCAGTACCTGGAGGATTCCTCA 190
QY 895 CAGTGTGTACAGTATACAGAGGATCTAGGACTTGAATTAAGAAGTCTACAATGACAG 954
DB 191 CCGAGGAACCTGTATCAGAGCAGGTTGGACTTCACACTTGACAAGCAGATTAACACAG 250
QY 955 CCCTTGGACAGCTGCTAAGATTACGTGTGATAAAGATAGCACAGTAAATGTTGAAGTT 1014
DB 251 TTCTAGGAACGGCTGCAAGGTTGCTTACAAAAGAGTCGACACAACTATGTTGATG 310
QY 1015 CAGGAAGTTCAGAGCTATTCCTAACCGTATTGCACTGATTAATTCGCAATTAAGAACAA 1074
DB 311 GCAGCACCCGAGAGAAGTGAATAGAGGTTGACAGATCAAAATCTCATTGAGGTAG 370
QY 1075 CAACCTTGACTTTGACCGTGAACAACTACAGAACCTTTGGCCAAATTAGCTGTTGGTG 1134
DB 371 CAGAGCAAGCTACGAGAGGAAAACCTCAATGAGAGGATTGCAAGCTCCCGGTTGGTG 430
QY 1135 TAGCTGTTATCAAGTAGGAGCTCCAACAGAGACAGCTTTTAAAGAAATGAACATTCGCA 1194
DB 431 TTGCTGTTATTCAGTGGGAGCACAACAGAACTGAACCTTAAGAGAGAAAGTTGGCAG 490
QY 1195 TTGAGGATGCTTAATGCTACAGCTGCAGCCGTTGAAGAAGGATTCGTTCTGCTGTTG 1254

DB 491 TTTAGGATGCTCTAAACGCAACCAAGGCTGCCGTTAGGAAGTATTCTTCTGTTGGTGA 550
QY 1255 GAACAGCAGCTATTACGGTTATTCAGAAAAGTAGCAGCTCTTTGAG 1298
DB 551 GGTGCATCTCTTTGAAGCTGGCTCTAAGGTTGATGCCATCAAG 594
RESULT 11
BE318972
LOCUS BE318972 695 bp mRNA EST 14-JUL-2000
DEFINITION NF043C12LF1F1087 Developing leaf Medicago truncatula cDNA clone
NF043C12LF 5', mRNA sequence.
ACCESSION BE318972
VERSION BE318972.1 GI:9192749
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
Fabales; Fabaceae; Papilionoideae; Medicago.
REFERENCE 1 (bases 1 to 695)
AUTHORS Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula leaf library
JOURNAL Unpublished (2000)
COMMENT Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7391
Fax: 580 221 7380
Email: gdmay@noble.org
Medicago Genome Initiative accession: MGI:S:27122
Insert length: 695 Std Error: 0.00
Plate: 043 row: C column: 12
Seq primer: TCACACAGAAACAGCTATGAC.
FEATURES
source
1..695
/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF043C12LF"
/clone_lib="Developing leaf"
/tissue_type="leaf"
/dev_stage="Pooled developmental"
/note="Vector: Lambda Zap; Contains a mixture of very
young, developing, mature and senescing leaves."
BASE COUNT 222 a 124 c 185 g 164 t
ORIGIN
Query Match 12.7%; Score 211.2; DB 106; Length 695;
Best Local Similarity 59.2%; Pred. No. 1.7e-48;
Matches 360; Conservative 0; Mismatches 248; Indels 0; Gaps 0;
QY 648 GACCTTGAAAACCCATTATCTTAATCAGGATAAAAAGTGTCAACATCCAAGACATT 707
DB 28 GAATTTGAGAACTCCAAGTTGCTTTTAGTGACAAAAGATATCCATGCAAGGATCTT 87
QY 708 TTGCCACTACTTGAGGAAGTTCTTTAAACCAACCGTCCATTACTTACTTATTCAGATGAT 767
DB 88 ATTAACATACTAGAGGACGCAATTAGAAACGGAATTCCTCTTTTGTGATTCATTGCAAGAT 147
QY 768 GTGATGCTGAGACAGCTTCACACCCCTTGTCTTGAACAGATTCGTTGACTTTCAATGTG 827
DB 148 ATTGAACAAGAAGCTTTGGCAACTCTTTGTTGTGAACAACACTTAGAGGATCACTGAAGAT 207
QY 828 GTTCTGTCAAAGCCGAGGATTTGGTATCGTCGTAAGCTATGCTTGAAGACATTGCT 887
DB 208 GCACACTTATAGGCTCTCGGATTTGGAGACGCAAGAGCCAACTACCTTGATGATTTGCC 267
QY 888 ATCTTGACAGTGCTACAGTGTATTACAGAGATCTAGGACTTGAATTTAAAGATGCTACA 947

```

Db 268 ATCTTGACAGGAGTACTGTATTACAGAGAGGTTGGCTTAGCTTAGACAAAGCTGGG 327
QY 948 ATGACAGCCCTTGACAGGCTGCTAGATTACAGTTGATAAAGATACACAGTAATTTGTT 1007
Db 328 ATGAGGTTCTAGGACTGCGGCCAGGTTGGTCTCACCAGAGATACACCAAGTAGTA 387
QY 1008 GAAGGTTTCAGGAAGTTGACAGAGCTATTGCTAACCGTATTGCACCTGATTAAATCGCAATTA 1067
Db 388 GGTGATGAAGTACCCCAAGAGCAGTTACCAACGGAGTTTCACAAATTAGAGAACCAAT 447
QY 1068 GAACAAACAACTTCTGACTTTGACCGTGAAACAACTCAAGAGCTTTGGGGAATTAGCT 1127
Db 448 GAGGCTGCAGAACAACTATGATGAAGAGGAGAGCTGAACGAAAGGATTGCAAACTGCT 507
QY 1128 GGTGGTGTAGCTGTTATCAAGATGAGAGCTCCCAACAGACAGCTTTAAAGAAATCAAA 1187
Db 508 GGTGGTGTGCTGTTATTCAGGTTGGTGACAAACTGAGACAGATCTCAAGGAAAGATA 567
QY 1188 CTTCCGATGAGGATGCTCTAAATGCTACACGTGCAGCGCTTGAAGAAAGTATCGTTGCT 1247
Db 568 TTGAGAGTTGAGGACGCTCTATATGCCACAAAGCGAGCTTTGAAAGGATTTGTAGTT 627
QY 1248 GGTGGTGG 1255
Db 628 GCGCGGGG 635

RESULT 12
BE131653
LOCUS
DEFINITION
L48-1652T3 Ice plant Lambda Uni-zap XR expression library, 48 hours
NaCl treatment Mesembryanthemum crystallinum cDNA clone L48-1652,
mRNA sequence.
ACCESSION
BE131653
VERSION
BE131653.1 GI:8579016
KEYWORDS
EST.
SOURCE
Mesembryanthemum crystallinum
common ice plant.
ORGANISM
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales;
Caryophyllales; Aizoaceae; Mesembryanthemum.
REFERENCE
1 (bases 1 to 710)
Cushman, J.C.
An expressed sequence tag database for the common ice plant,
Mesembryanthemum crystallinum
Unpublished (1997).
JOURNAL
Contact: Cushman JC
Department of Biochemistry and Molecular Biology
Oklahoma State University
350 Noble Research Center, Stillwater, OK 74078-3035, USA
Tel: 405-744-6207
Fax: 405-744-7799
Email: jcushman@biochem.okstate.edu
PCR Primers
FORWARD: T7
BACKWARD: T3
Plate: L48-17 row: E column: 4
Seq primer: T3
High quality sequence stop: 350
POLYA-No.
Location/Qualifiers
1. .710
/organism="Mesembryanthemum crystallinum"
/db_xref="taxon:3544"
/clone="L48-1652"
/clone_lib="Ice plant Lambda Uni-zap XR expression library
, 48 hours NaCl treatment"
/tissue_type="Leaf, 48 h 0.4M NaCl"
/dev_stage="Six week old"
/note="Vector: Lambda Uni-zap XR, Bluescript SK-; Site_1:
EcoRI; Site_2: XhoI"
BASE COUNT 225 a 122 c 194 g 169 t
```

ORIGIN

```

Query Match 12.7%; Score 210.8; DB 104; Length 710;
Best Local Similarity 58.0%; Pred. No. 2.2e-48;
Matches 391; Conservative 0; Mismatches 282; Indels 1; Gaps 1;

QY 582 GGATGCAATTTGACCGTGGTTACCTGTCTCAATACATGCTACAGACATGAAAAATG 641
Db 1 GGAATGCAATTTGATCGAGCTATATCTCCCATATCTTTGTACACATAGTATAAATG 60
QY 642 GTTGACAGCTTTGAAACCCCATTTATCTTAATCAGCGTAAAAAGTGTCAACATCCAA 701
Db 61 ATTGTTGAATATCAAAACTGTAAAGTTGCTTGGTAGACAGAAAGATCACAACTGCGGA 120
QY 702 GACATTTTCCCACTACTTTCAGGAGTTCCTTAACACCAACCGTCCATTAATTTGCA 761
Db 121 GACCTAGTTAAATGCTTAGAAGATGCGATTAAACAACTACCCAACTTTGATTAATCGCA 180
QY 762 GATGATGTGGATGGTGAAGCAGCTTCCAAACCCCTTGTCTTGAACAAGATTCGTGTACTTTC 821
Db 181 GAAGACATTTGACGAGCAACCTCTTGTCTACATTAAGTTGTAACAAGCTCCGGGATCACATT 240
QY 822 AATGCTGTTGCTGTCAAAGCGCAGGATTTGGTATGCTGCTGTAAGAGCTATGCTTGAAGAC 881
Db 241 AGGATAGCTGCAATGAAGGCTCCGGCTTATGGAGATCGAAGCATCAATATCTTTGATGAT 300
QY 882 ATTGCTATCTTGACAGGTGCTACAGTGATTACAGAGATCTAGGACTTGAATTTAAAGAT 941
Db 301 ATTGCTATCTTACTGAGGAGACTGTGATCAGAGAGAGGTTGGTTTGTCTTTTGGACAAA 360
QY 942 GGTACAATGACAGCCCTTGGACAGGCTGTCTGAAGATTACAGTTGATTAAGATAGACACAGTA 1001
Db 361 GTTGGCACAGAGGCTCTTGGGACATGCTTCAAGGCTGCTCAAAAAGAGTCAAGTACC 420
QY 1002 ATTGTTGAAGTTTCAGGAAGTTTCAGAAAGTATTTGCTTAACCGTATTGCACTGATTAATCG 1061
Db 421 ATTGTTGGGGTGGGAGCACACAAGAACGAGTTACTAGGAAATTTGCCAGATTAGGAAG 480
QY 1062 CAATTAGAAACAACAACCTTCTGACCTTGACCGTGAACAACTACAAAGACGTTTGGCGAAA 1121
Db 481 CAGCTGAGGATGCTGACCCAGGAGTACGAAAAACAAGCTCAATGAAGGATAGGACAAA 540
QY 1122 TTAGCTGGTGGTGTAGCTGTTATCAAAAGTAGGAGCTCCACAGAGACAGCTTTAAAGAA 1181
Db 541 CTCTCCAGCGGTGCTGCTGATTCAGGTTGGAGCCAGACTGAGACTGAGATGACAGAA 600
QY 1182 ATGAACCTTCGCATTTGAGGATGCTCTTAATGCTACAGTGCAGCCGCTTGAAGAGGTATC 1241
Db 601 AAGAACTGAGAGTGGAGGATGCTCTTAATGC-ACAAAGGCAGCTGTAGATGAAGGTATC 659
QY 1242 GTTGCTGGTGGTGG 1255
Db 660 GGTGATGAGTGGAGG 673
```

RESULT 13

AW216817

LOCUS

DEFINITION

AW216817

VERSION

KEYWORDS

SOURCE

ORGANISM

Lycopersicon esculentum

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE

AUTHORS

Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E., Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Alm,S., Ronning ,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.

TITLE Generation of ESTs from tomato callus tissue
 UNPUBLISHED (1999)
 COMMENT Contact: David Frisch
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 4366
 Fax: 864 656 4293
 Email: dfrisch@CLEMSON.EDU
 5 prime sequence.

FEATURES	SOURCE
----------	--------

```

location/ydairies
1. .612
/organism="Lycopersicon esculentum"
/cultivar="TA96"
/db_xref="taxon:4081"
/clone="c1EC87D23"
/clone_lib="tomato callus, TAMU"
/tissue_type="callus"
/dev_stage="25-40 days old"
/lab_host="X11-Blue MRF"
/note="vector: pBluescript SK(-); S
xhoI; supplier: Giovannoni laborator
of seedlings 7-10 days post-germina
at both ends and placed on MS media
Mixed callus was harvested at 25 an
undifferentiated masses. Tomato call
95 c 160 g 166 t 1
190 a
BASE COUNT

```

Query Match	12.6%	Score 209.4;	DB 87;	Length 612;
Best Local Similarity	58.8%;	Pred. No. 5.1e-48;		
Matches 360;	Conservative	0;	Mismatches 252;	Indels 0;
				Gaps 0;

513	QY	GTGGCAGCATGGTGTGATTACCATCAAGAATCTCGAGGTATGGAACAGAACTTGAA	572
1	Db	CTTGAGGAAGGTGTGTGCACACTAGAGAGGGTAAAGTGTGTAACACAGTCTGGCT	60
573	QY	GTGGTTGAAGCATCTCAATTTGACCGTGGTTACCTGTCTCAATACATGGTCCACAGCAAT	632
61	Db	GTGGTTGAAGGAATGCAATTTGACCGTGGTTATGTCCTCTCTACTGTTTGTACCACAGT	120
633	QY	GAATAAATGGTTGCAGACCTTGAAACCCATTTATCTTAATCACGGATAAAAAAGTGCA	692
121	Db	CAGAAATGTCGGTGAATATGAGAACTGAAGTTGCTACTGGTTGATAAAGAATAACA	180
593	QY	AACATCCAAGACATTTTGCCACTACTTGAGGAAGTCTTAAACCAACCGTCCATTACTC	752
181	Db	AATCAAGAGATCTGTGTTAATGTCTGGAAGATGCTATCAGAAATGGTTACCCAAATTTA	240
753	QY	ATTATTCCAGATGATGGATGGTCAAGCACTTCCAACCCCTTGCTTCAACAAGATTTGCT	812
241	Db	ATTATGCTGAAGATATTGACGAGAAGCTTGGCAACTCTGTTGTCAATAAGCTTAGA	300
813	QY	GGTACTTCAATGGTTGTGCTCAAAAGCGCAGGATTTGGTGTGCTGCTGAAGCTATG	872
301	Db	GGTGGCTGAAGTCGCTGCACCTTAAAGCTCTGGTTTTGGTGCGGAAAGACCAATAT	360
873	QY	CTTGAGACATTTGCTATCTTGACAGTGGTACAGTGATTACAGAGGATCTAGGACTTGAA	932
361	Db	CTTGATGACATPAGCAACCCCTTACTGGAGGCACCTGTTATTAGGAGGAGCTTGGCCTTACC	420
933	QY	TTAAGAATGCTACAATGACACGCCCTTGGACAGGCTGCTAGAGTTACAGTTGATTAAGAT	992
421	Db	TTGACAGAGGCTGACAGGAAGTTCTTAGGTGATGCTGCTAAAGTAGTCTGACTAAGGAT	480
993	QY	AGCAGAGTAATTTGTGAAGGTTTCAGAGGATTTGCTTAACCGTATTTCACCTG	1052
481	Db	GCCACTACAATTTGGTGTGCTGATAGCACTCAGGAAGCAGTCAACANACGTGTTGCCACAG	540
1053	QY	ATTAAATCGCAATTTAGAACAACAACCTTCTGACTTTTGCCGCTGAAAAACTACAAGAAGCT	1112
541	Db	ATTAAAAACCTTGATAGAGGCTGCGAGTCAAGATTATGAAGAAGGAAGAGCTAAATCAAGA	600

Oy	1113	TTGGCGAAATTA 1124 I I I I I
Db	601	ATTGCTAAGTTA 612
RESULT	14	
BE231179		
LOCUS		902 bp mRNA EST 25-JUL-2000
DEFINITION		HVSMEG0017G14f Hordeum vulgare pre-anthesis spike EST library HVCNDA0008 (white to yellow anther) Hordeum vulgare cDNA clone
VERSION		HVSMEG0017G14f, mRNA sequence.
KEYWORDS		BE231179
SOURCE		BE231179.1 GI:8967402
ORGANISM		EST. barley. Hordeum vulgare
REFERENCE		Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Hordeum.
AUTHORS		1 (bases 1 to 902) Wing,R., Close,F.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D., Yu, ,Y., Anderson,H., Dale,J., Henry,D., Kernodle,S., Palmer,M., Rambo ,T., Saski,C., Schwartzbeck,J., Simmons,J., Choi,D.W., Main,D. and wood,T.
TITLE		Development of a genetically and physically anchored EST resource for barley genomics
JOURNAL		Unpublished (2000)
COMMENT		Contact: Wing RA Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 7288 Fax: 864 656 4293 Email: rwing@clemson.edu Seq primer: AATTACCCCTCACTAAAGGG High quality sequence start: 7 High quality sequence stop: 578. Location/Qualifiers 1..902 /organism="Hordeum vulgare" /cultivar="Morex" /db_xref="taxon:4513" /clone="HVSMEG0017G14f" /clone_lib="Hordeum vulgare pre-anthesis spike EST library HVCNDA0008 (white to yellow anther)" /tissue_type="pre-anthesis spike" /lab_host="SOLR" /note="vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI"
FEATURES		
SOURCE		
BASE COUNT	256 a	199 c 250 q 196 t 1 others .

	Query Match	12.6%	Score 208.8;	DB 105;	Length 902;
	Best Local Similarity	57.6%;	Pred. No. 8.8e-48;		
	Matches 392;	Conservative 0;	Mismatches 288;	Indels 1;	Gaps 1;
Qy	587	GCAATTTGACCGTGGTACCTGCTCTCAATACATCGTCCACACAAATGAAATAATGGTTGC	646		
Db	9	GGAGATTGACCGTGGCTACATTTCCCTCAATTCTGTCACAAACCTTCGAGAATCGATGT	68		
Qy	647	AGACCTTGAAAACCCATTTATCTTTAATCACGGATATAAAAAGTGTCAACATCTCCAGACAT	706		
Db	69	CGAGTTTGAGAAATGCTAGAGTCTTATAACTGTATCAGAAGATCAACAAGCATAAAGGAAAT	128		
Qy	707	TTTGCCACTACTTGAGGAAGTCTTTAAACCAACCGTCTTACTTCTCATTTATTTGCAGATGA	766		
Db	129	CATTCCACTTCTGGAGCAGACATACACAGTTGAGATGTCCTGATTCATTGTAGCCGAGGA	188		
Qy	767	TGTGGATGGTGAAGACTTCCAAACCTTGTCTTGAAACAAGATTCGTGGTACTTTCAATGT	826		
Db	189	CATTACTGGTGAAGCTTTGGCGACTCTGTGTGTTAAACAAGCTCAGAGGTATTATTAACGT	248		
Qy	827	GGTTTGCTCTAAAGCGCCAGGATTTTGGTGTATCGTCGTAAGACTATGCTTGAAGACATTCG	886		


```

Db 249 TCGCGCATCAAAAGCCCAAGTTTCGGTGAGCGCGGAAGGCTGTCTTCAGGATATTC 308
QY 887 TATCTTGACAGGTGGTACAGTATACAGAGGATCTAGGACTTGAATTAAGATGCTAC 946
Db 309 CATCGTGACAGGTGCTGAATACCTACGAGGATCTTGGTCTGGTTGAGATCCGAC 368
QY 947 AATGACAGCCCTTGGACAGGCTGCTGAATACAGTATGATAAGATGACAGTAAATTC 1006
Db 369 GGTAGACCAACTTGGACAGGAGGAAATCAATTCATCAGACTACAACCCCTCAT 428
QY 1007 TGAAGCTTCAGGAAGTTCAGAAGCTATTGCTAACCGTATTGCACTGATTAATCGCAAT 1066
Db 429 AGCAGATGACAGTACGAGGAGGATCCAGCGAGGGTTGACAGCTAAGAGAGGAGCT 488
QY 1067 AGAAACAAACACTTCTGACTTTGACCGTGAAGAACTACAGAACTGTTGGCGAAATTAGC 1126
Db 489 TTCTGAACTGATTCATATGATCTCAGAAATTTGGCTGAGAGATTTGCCAAGCTTTC 548
QY 1127 TGGTGTGTAGCTGTTTCAAGTAGGAGCTCCACAGACAGAGCTTTTAAAGAAATGAA 1186
Db 549 TGGTGTGTGGCGTCAATCAAGGTTGGAGNACAACTGAGACAGAGCTCCGAGACCGTCA 608
QY 1187 ACTTCCATTGAGGATGCTCTAAATCTACAGCTGACGCGCTTGAAGAAGGTATCGTTCC 1246
Db 609 GCTGCGGATCGAGGAGCGAAGACCCACTTTCGCTG-CATCGAGGAAGGCAATTTGGTCC 667
QY 1247 TGGTGTGTGGAACAGCACTTAT 1267
Db 668 CGGTGTGTGGCGGATATGT 688

```

```

RESULT 15
AW622566 739 bp mRNA EST 28-MAR-2000
LOCUS ESR313366 tomato root during/after fruit set, Cornell University
DEFINITION Lycopersicon esculentum cDNA clone cLEX15J21 5', mRNA sequence.
ACCESSION AW622566
VERSION AW622566.1 GI:7334213
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids
I; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 739)
AUTHORS van der Hoeven,R.S., Garvin,D., Matern,A.L., Holt,I.E., Kochian,L.,
Liang,F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley
,S.D.
Generation of ESTs from tomato root, during and after fruit set
Unpublished (1999)
Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.
FEATURES
Location/Qualifiers
1..739
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEX15J21"
/clone_lib="tomato root during/after fruit set, Cornell
University"
/tissue_type="root"
/dev_stages="plants during and after fruit-set"
/note="Vector: pBluescript SK(-); Site:1: EcoRI; Site:2:
XhoI; supplier: Tanksley; Tissue supplied by Dave Garvin
(USDA-ARS, Ithaca, NY 14850)."
```

```

BASE COUNT 234 a 125 c 200 g 179 t 1 others
ORIGIN
Query Match 12.5%; Score 207.4; DB 93; Length 739;
Best Local Similarity 58.2%; Pred. No. 2e-47;
Matches 383; Conservative 0; Mismatches 272; Indels 3; Gaps 1;
QY 13 ATATGCAAAAGAAATCAAAATTTTCAGCAGATGCGGGTCTGCGATGTCGCGGAGTTG 72
Db 78 ATGCTGCTAAGGATATAATTTGGAGTTGAAGTAGGGAATATGCTTCAAGAGTTG 137
QY 73 ATATGTTACGATACCGTCAAGTAACGCTTGGTCTTAAAGGGCGCAATGTGTCTTTC 132
Db 138 AGCAGCTTCTGCTGCGGATGTTGACCAATGTTGTCGCTGCTGCTGCTGCTGCTGCTG 197
QY 133 AAAAAGCTTTTGGTCTCCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 192
Db 198 ACAAAGTTGGGGTGCACCCCAAGGTGACAAAAGATGGTGTCTACTGTTGCAAAAAGCAT 257
QY 193 AATTAGAAGATCATTTTGAAGAACATGGGAGCAAAATTTGGTCTGTAAGTGGCTTCTAAA 252
Db 258 AATTCAAGGATAAATAAATAATTTGGTCCAGCCTTTGAAGCAAGTTGCCAATGCCA 317
QY 253 CCAATCATATTGCTGGTATGGGACGACTACTGCAACAGTTTTTGACACAAGCCATTGTT 312
Db 318 CTAATGATGCTGCTGGGATGTTGACCACTTTGTGCAACAGTACTCACCCGAGCAATATT 377
QY 313 ATGAAGGACTAAAATGTGACAGGAGGTGCTTAATCCAAATTTGGTATCCGTCGAGGCA 372
Db 378 CTGAAGGGTGCAGTCTGTAGCTGTGGTATGAATGCAATGCAATGCAATGCAATGCAAT 437
QY 373 AAACAGCAACAGCAGCTGTTGAAGCCTTCAAGCCCTTCAAGCCCTTCAAGCCCTTCAAG 432
Db 438 CAATGGCTGCTGATGCTGTTGTAACAACCTGAAAGCAAGCAGCAGGATGATAGCAAT 497
QY 433 AGGAAGCTATTGCTCAGGTCGCTGCAAGTATATCATCGCTCTGAA---AAAGTTGGAGAT 489
Db 498 CCGAGGAGATTGCCAGGTTGGTACAATCTCTGCAAAATGGAGAACGGAATAGGTGAGA 557
QY 490 ATATCTCAGAACTATGAGCGTGTGGGCAACGATGGTGTGATTACCATCGAAGATCTC 549
Db 558 TAAATGCAAGGCTATGGAGCAAGTAGGAAAGAGGTGTCTACTAATTTCAAGATGGAA 617
QY 550 GAGGTATGAAACAGAACTTTGAAGTGGTTGAAGGATGCAATTTGACCGTGTACCTGT 609
Db 618 AGACACTGCTTAATGATTTGCAAGTTGTTGAGGGGATGAAGTTGACAGGGGCTACAT 677
QY 610 CTCAATACATGTCACAGCAATGAAAATTTGGTTCAGACCTTGAAGAACCCATTAT 667
Db 678 CCCCATACTTCATCANATGAGAAGAAATCAGAAAATGTGAACCTGGATGACCCACTTAT 735

```

Search completed: February 16, 2001, 01:02:10
Job time: 10917 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2001, 20:33:17 ; Search time 1836.42 Seconds
(without alignments)
3900.644 Million cell updates/sec

Title: US-09-001-737-7

Perfect score: 1661

Sequence: 1 GAATTCGGCTTCATATGGCA.....TGGCGGATAAGCCGAATTC 1661

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 12362702 seqs, 2156297445 residues

Total number of hits satisfying chosen parameters: 24725404

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents_NA.*
1: /cgn2_6/ptodata/2/pna/pctus_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US080_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US081_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US083_COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US084_COMB.seq.*
9: /cgn2_6/ptodata/2/pna/US085_COMB.seq.*
10: /cgn2_6/ptodata/2/pna/US086_COMB.seq.*
11: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
12: /cgn2_6/ptodata/2/pna/US088_COMB.seq.*
13: /cgn2_6/ptodata/2/pna/US089_COMB.seq.*
14: /cgn2_6/ptodata/2/pna/US090_COMB.seq.*
15: /cgn2_6/ptodata/2/pna/US091_COMB.seq.*
16: /cgn2_6/ptodata/2/pna/US092_COMB.seq.*
17: /cgn2_6/ptodata/2/pna/US093_COMB.seq.*
18: /cgn2_6/ptodata/2/pna/US094_COMB.seq.*
19: /cgn2_6/ptodata/2/pna/US095A_COMB.seq.*
20: /cgn2_6/ptodata/2/pna/US095B_COMB.seq.*
21: /cgn2_6/ptodata/2/pna/US095C_COMB.seq.*
22: /cgn2_6/ptodata/2/pna/US096A_COMB.seq.*
23: /cgn2_6/ptodata/2/pna/US096B_COMB.seq.*
24: /cgn2_6/ptodata/2/pna/US097_COMB.seq.*
25: /cgn2_6/ptodata/2/pna/US098_COMB.seq.*
26: /cgn2_6/ptodata/2/pna/US099_COMB.seq.*
27: /cgn2_6/ptodata/2/pna/US099A_COMB.seq.*
28: /cgn2_6/ptodata/2/pna/US099B_COMB.seq.*
29: /cgn2_6/ptodata/2/pna/US099C_COMB.seq.*
30: /cgn2_6/ptodata/2/pna/US099D_COMB.seq.*
31: /cgn2_6/ptodata/2/pna/US099E_COMB.seq.*
32: /cgn2_6/ptodata/2/pna/US099F_COMB.seq.*
33: /cgn2_6/ptodata/2/pna/US099G_COMB.seq.*
34: /cgn2_6/ptodata/2/pna/US099H_COMB.seq.*
35: /cgn2_6/ptodata/2/pna/US099I_COMB.seq.*
36: /cgn2_6/ptodata/2/pna/US099J_COMB.seq.*
37: /cgn2_6/ptodata/2/pna/US099K_COMB.seq.*
38: /cgn2_6/ptodata/2/pna/US099L_COMB.seq.*
39: /cgn2_6/ptodata/2/pna/US099M_COMB.seq.*
40: /cgn2_6/ptodata/2/pna/US099N_COMB.seq.*
41: /cgn2_6/ptodata/2/pna/US099O_COMB.seq.*
42: /cgn2_6/ptodata/2/pna/US099P_COMB.seq.*
43: /cgn2_6/ptodata/2/pna/US099Q_COMB.seq.*

44: /cgn2_6/ptodata/2/pna/US6019_COMB.seq.*
45: /cgn2_6/ptodata/2/pna/US6020_COMB.seq.*
46: /cgn2_6/ptodata/2/pna/US6021_COMB.seq.*
47: /cgn2_6/ptodata/2/pna/US6022_COMB.seq.*
48: /cgn2_6/ptodata/2/pna/US6023_COMB.seq.*
49: /cgn2_6/ptodata/2/pna/US6024_COMB.seq.*
50: /cgn2_6/ptodata/2/pna/US6025_COMB.seq.*
51: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq.*
52: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq.*
53: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq.*
54: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq.*
55: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
56: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1661	100.0	1661	14	US-09-001-737-7 Sequence 7, Appli
2	1052	63.3	1654	14	US-09-001-737-3 Sequence 3, Appli
3	1039.2	62.6	1623	21	US-09-583-110-1024 Sequence 1024, Ap
4	1039.2	62.6	1623	42	US-60-174-089-100 Sequence 100, App
5	1038.4	62.5	1926	22	US-09-613-303-50 Sequence 50, Appli
6	1035.4	62.3	5365	13	US-08-961-527-77 Sequence 77, Appli
7	1024.4	61.7	3714	31	US-60-061-998-525 Sequence 525, App
8	1023.4	61.6	5373	27	US-60-029-960-223 Sequence 223, App
9	1016.6	61.2	2597	31	US-60-068-175-420 Sequence 420, App
10	945.8	56.9	8787	31	US-60-068-217-910 Sequence 910, App
11	945.8	56.9	8784	29	US-60-045-649-1072 Sequence 1072, Ap
12	945.2	56.9	8784	29	US-60-046-653-1063 Sequence 1063, Ap
13	942.6	56.7	3625	14	US-09-070-927-42 Sequence 42, Appli
14	942.6	56.7	3625	14	US-09-070-927A-42 Sequence 42, Appli
15	934	56.2	1638	15	US-09-134-000-1154 Sequence 1154, Ap
16	901.6	54.3	2237	31	US-60-068-186-529 Sequence 529, App
17	901.6	54.3	2300	30	US-60-050-444-525 Sequence 525, App
18	741.4	44.6	53996	12	US-08-831-156A-95 Sequence 95, Appli
19	741.4	44.6	278942	22	US-09-611-529-7437 Sequence 7437, Ap
20	735.6	44.3	1623	18	US-09-450-969-2422 Sequence 2422, Ap
21	704.2	42.4	1647	14	US-09-077-574A-1 Sequence 1, Appli
22	704.2	42.4	4972	14	US-09-077-574A-5 Sequence 5, Appli
23	701.6	42.2	6418	28	US-60-038-081-257 Sequence 257, App
24	701.6	42.2	6418	29	US-60-046-933-256 Sequence 256, App
25	701.6	42.2	6418	31	US-60-068-228-260 Sequence 260, App
26	657.4	39.6	1641	20	US-09-540-236-654 Sequence 654, App
27	657.4	39.6	1641	37	US-60-128-476-1558 Sequence 1558, Ap
28	654.2	39.4	96109	39	US-09-596-002-35 Sequence 35, Appli
29	654.2	39.4	96109	39	US-60-140-121-35 Sequence 35, Appli
30	652	39.3	1368	22	US-09-611-529-119 Sequence 119, App
31	649.6	39.1	24380	31	US-60-068-139-393 Sequence 393, App
32	649.6	39.1	24436	33	US-60-082-116-300 Sequence 300, App
33	624.8	37.6	1653	17	US-09-328-352-1382 Sequence 1382, Ap
34	606	36.5	1050	15	US-09-107-532-3030 Sequence 3030, Ap
35	590.6	35.6	1665	14	US-09-001-737-1 Sequence 1, Appli
36	589.8	35.5	1647	18	US-09-472-971-5 Sequence 5, Appli
37	589.8	35.5	4524	18	US-09-472-971-7 Sequence 7, Appli
38	572.8	34.5	1659	16	US-09-252-691-1518 Sequence 1518, Ap
39	572.8	34.5	1659	16	US-09-252-691C-1518 Sequence 1518, Ap
40	569.6	34.3	1674	16	US-09-252-691-1525 Sequence 1525, Ap
41	569.6	34.3	1674	16	US-09-252-691C-1525 Sequence 1525, Ap
42	568.6	34.2	966	55	US-09-107-433-1679 Sequence 1679, Ap
43	568.2	34.2	1662	14	US-09-001-737-5 Sequence 5, Appli
44	551.6	33.2	910715	1	PCT-US98-12764-1 Sequence 1, Appli
45	544.4	32.8	2416	16	US-09-221-017B-895 Sequence 895, App

ALIGNMENTS

RESULT 1
 US-09-001-737-7
 : Sequence 7, Application US/09001737
 : GENERAL INFORMATION:
 : APPLICANT: Mizzen, Lee
 : APPLICANT: Wisniewski, Jan
 : TITLE OF INVENTION: STREPTOCOCCAL HEAT SHOCK PROTEINS OF THE
 : TITLE OF INVENTION: HSP60 FAMILY
 : NUMBER OF SEQUENCES: 91
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: SEED and BERRY LLP
 : STREET: 701 Fifth Avenue, 6300 Columbia Center
 : CITY: Seattle
 : STATE: Washington
 : COUNTRY: USA
 : ZIP: 98104
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/001,737
 : FILING DATE: 31-DEC-1997
 : CLASSIFICATION: 435
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Sharkey, Richard G.
 : REGISTRATION NUMBER: 32,629
 : REFERENCE/DOCKET NUMBER: 870109.408
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (206) 622-4900
 : TELEFAX: (206) 682-6031
 : INFORMATION FOR SEQ ID NO: 7:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 1661 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: 15..1649
 : US-09-001-737-7

Query Match 100.0%; Score 1661; DB 14; Length 1661;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1661; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAATTCGGCTTCATATGGCAAAAGAAATCAAAATTTTCAGCAGATGCGCGTGGCCATGG 60
 Db 1 GAATTCGGCTTCATATGGCAAAAGAAATCAAAATTTTCAGCAGATGCGCGTGGCCATGG 60
 QY 61 TCGCGGAGCTTCATATGTTAGCAGATACCGTCAAAAGTAAAGCTTGGTCTCAAGGCGCA 120
 Db 61 TCGCGGAGCTTCATATGTTAGCAGATACCGTCAAAAGTAAAGCTTGGTCTCAAGGCGCA 120
 QY 121 ATGTTGTTCTTGAAGAAAGCTTTTGGTTCTCCCTTAATTAATACGCGGGTAAACCATTTG 180
 Db 121 ATGTTGTTCTTGAAGAAAGCTTTTGGTTCTCCCTTAATTAATACGCGGGTAAACCATTTG 180
 QY 181 CTAAGAGATCGAATYTAGAAGATCAATTTTGAACACATGGGAGCAAAATTTGGTGTCTGAAG 240
 Db 181 CTAAGAGATCGAATYTAGAAGATCAATTTTGAACACATGGGAGCAAAATTTGGTGTCTGAAG 240
 QY 241 TGGCTTCTAAAACCAATGATATTGCTGGTATGGGACGACTACTCAACAGTTTTCACAC 300
 Db 241 TGGCTTCTAAAACCAATGATATTGCTGGTATGGGACGACTACTCAACAGTTTTCACAC 300
 QY 301 AAGCCATTGTTCAAGGACTAAAAAATGTGACAGCAGGTGCTAATCCAATTTGGTATCC 360
 Db 301 AAGCCATTGTTCAAGGACTAAAAAATGTGACAGCAGGTGCTAATCCAATTTGGTATCC 360
 QY 361 GTCGAGGCAATTTGAACAGCAGCAACAGCTGTTGAAGCCTTGAAGCCATTGCTCAAC 420

Db 361 GTCGAGGCAATTTGAACAGCAGCAACAGCTGTTGAAGCCTTGAAGCCATTGCTCAAC 420
 QY 421 CTGTATCTGCAAGGAAGCTATTGCTCAGTCCGCTCAGTATCATCACGCTCTGAAAGAG 480
 Db 421 CTGTATCTGCAAGGAAGCTATTGCTCAGTCCGCTCAGTATCATCACGCTCTGAAAGAG 480
 QY 481 TTGGAGAGTATATCTCAGAAAGCTATGGAGCGTGTGGGCAACGATGTTGATATACCATCG 540
 Db 481 TTGGAGAGTATATCTCAGAAAGCTATGGAGCGTGTGGGCAACGATGTTGATATACCATCG 540
 QY 541 AGAATCTCGAGGTATGGAAACAGAACTTTGAAGTGGTTGAAGCATGCAATTTGACCGTG 600
 Db 541 AGAATCTCGAGGTATGGAAACAGAACTTTGAAGTGGTTGAAGCATGCAATTTGACCGTG 600
 QY 601 GTTACCTGCTCAATACATGTTGACAGCAATGAAAAAATGTTGAGACCTTGAAGACC 660
 Db 601 GTTACCTGCTCAATACATGTTGACAGCAATGAAAAAATGTTGAGACCTTGAAGACC 660
 QY 661 CATTTATCTTAATACGCGATAAAAAGTGTCAAAACATCCAAGACATTTTGCACACTACTTG 720
 Db 661 CATTTATCTTAATACGCGATAAAAAGTGTCAAAACATCCAAGACATTTTGCACACTACTTG 720
 QY 721 AGAAGTCTTAAACCAACCGTCCATTACTCATTTATTCAGATGATGGATGGTGAAG 780
 Db 721 AGAAGTCTTAAACCAACCGTCCATTACTCATTTATTCAGATGATGGATGGTGAAG 780
 QY 781 CACTTCCAAACCTTGTCTTGAACAGATTCGTTGATGTTCAATGCTGTTGCTGTCAGAG 840
 Db 781 CACTTCCAAACCTTGTCTTGAACAGATTCGTTGATGTTCAATGCTGTTGCTGTCAGAG 840
 QY 841 CGCCAGGATTTGGTATCGTCTGTAAGCTATGCTTGAAGACATTTGATCTTGCACAGGTG 900
 Db 841 CGCCAGGATTTGGTATCGTCTGTAAGCTATGCTTGAAGACATTTGATCTTGCACAGGTG 900
 QY 901 GTACAGTGAATACAGAGATCTAGGACTTTGAATTTAAAGATGCTACAATGACAGCCCTTG 960
 Db 901 GTACAGTGAATACAGAGATCTAGGACTTTGAATTTAAAGATGCTACAATGACAGCCCTTG 960
 QY 961 GACAGGCTCTAAGATTACAGTTGATAAGATAGACAGCTAATTTGTTGAAGTTTCAGGAA 1020
 Db 961 GACAGGCTCTAAGATTACAGTTGATAAGATAGACAGCTAATTTGTTGAAGTTTCAGGAA 1020
 QY 1021 GTTCAGAACTATTGCTAACCGTATTGCTACTGATTAATTAATGCAATTTAGAAACAACTT 1080
 Db 1021 GTTCAGAACTATTGCTAACCGTATTGCTACTGATTAATTAATGCAATTTAGAAACAACTT 1080
 QY 1081 CTGACTTTGACCGTGAAAAAAGCTTACAGAAAGCTTTTGGCGAAATTTAGCTGGTGTAGCTG 1140
 Db 1081 CTGACTTTGACCGTGAAAAAAGCTTACAGAAAGCTTTTGGCGAAATTTAGCTGGTGTAGCTG 1140
 QY 1141 TTATCAAAAGTAGGAGCTCCACAGACAGCTTTTAAAGAAATGAACTTCCCATTTGAGG 1200
 Db 1141 TTATCAAAAGTAGGAGCTCCACAGACAGCTTTTAAAGAAATGAACTTCCCATTTGAGG 1200
 QY 1201 ATGCTCTAAATGCTACACGTCAGCGCTTGAAGAGAGGTATCGTTGGTGGTGAAGACAG 1260
 Db 1201 ATGCTCTAAATGCTACACGTCAGCGCTTGAAGAGAGGTATCGTTGGTGGTGAAGACAG 1260
 QY 1261 CACTTATACGGTTATTGAAAAAGTAGCAGCTCTTGAGCTTGAGGCGGATGATGCTACTG 1320
 Db 1261 CACTTATACGGTTATTGAAAAAGTAGCAGCTCTTGAGCTTGAGGCGGATGATGCTACTG 1320
 QY 1321 GACGTAACATTTGCTGCTCTAGAGAGCGCTCTAGCTCAAAATGCTTTAAATGCTG 1380
 Db 1321 GACGTAACATTTGCTGCTCTAGAGAGCGCTCTAGCTCAAAATGCTTTAAATGCTG 1380
 QY 1381 GGTACGAAGGCTCCCTAGTATTGACAGTGTGAAAAAAGCAGCTGCGAGGACAGGATTTA 1440
 Db 1381 GGTACGAAGGCTCCCTAGTATTGACAGTGTGAAAAAAGCAGCTGCGAGGACAGGATTTA 1440
 QY 1441 ATGCTGCAACAGGCTGAGTGGGTTGATGATTTAAACAGGAATCATTTGACCTGTCAGAG 1500

Db 1441 ATGCTGCAACAGGTGAGTGGTGTGATATGATTAACAAACAGGAATCATTTGACCCCTGTCAAAG 1500
QY 1501 TAACACGATCAGCGCTTCAAAATGACAGCTCTGTAGTAGTCTTATTTTGAACACAGAAAG 1560
Db 1501 TAACACGATCAGCGCTTCAAAATGACAGCTCTGTAGTAGTCTTATTTTGAACACAGAAAG 1560
QY 1561 CAGTGTGCTTAATAAACCTGACACAGCTACGCCAGCGCCAGCAATGCCAGCAGGTATGG 1620
Db 1561 CAGTGTGCTTAATAAACCTGACACAGCTACGCCAGCGCCAGCAATGCCAGCAGGTATGG 1620
QY 1621 ATCCAGGAATGATGGGTGGGATGGCGGATGAAGCCGAATTC 1661
Db 1621 ATCCAGGAATGATGGGTGGGATGGCGGATGAAGCCGAATTC 1661

RESULT 2

US-09-001-737-3
; Sequence 3, Application US/09001737
; GENERAL INFORMATION:
; APPLICANT: Mizzen, Lee
; APPLICANT: Wisniewski, Jan
; TITLE OF INVENTION: STREPTOCOCAL HEAT SHOCK PROTEINS OF THE
; TITLE OF INVENTION: HSP60 FAMILY
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 701 Fifth Avenue, 6300 Columbia Center
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/001,737
; FILING DATE: 31-DEC-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 870109.408
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1654 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 15..1637
US-09-001-737-3

Query Match 53.3%; Score 1052; DB 14; Length 1654;
Best Local Similarity 77.7%; Pred. No. 3e-273;
Matches 1271; Conservative 0; Mismatches 365; Indels 0; Gaps 0;
QY 1 GAATTCGGCTTCATATGGCAAAAGAAATCAAAATTTTCAGCAGATCGCGCTGTGCCATGG 60
Db 1 GAATTCGGCTTCATATGGCAAAAGAAATTAATTTTCATCAGATGCCGCTTCAGCTATGG 60
QY 61 TGGCCGGAGTTGATATTTAGCAGATACCGTCAAGTAAAGCTTTGGTCTTAAAGGGCCCA 120
Db 61 TCCGTGTGTCGATATCCTTGCAGACTGTAAAGTAACTTTGGGACCAAAAGGTGCGCA 120
QY 121 ATGTTGTTCTTGAAGAGCTTTTGGTCTCCCTTAATTAATGACGGGTAAACCAATTG 180
Db 121 ATGCTCTAAACGCTTACTCGTGCAGCTGTTGAAGAGGTATTTGTTGCAGGTGGTGAACAG 1260

Db 121 ATGTCGTTCTTGAAGAGCTCATTCGGTTACCCCTTGATTTACCAATGACGGTGTGACTATTG 180
QY 181 CTAAGAGATCGAATTAGAAAGATCAATTTTGAACACATGGGAGCAAAATTTGCTGTCTGAAG 240
Db 181 CCAAGAAATTTGAATTAGAAAGACCATTTTGAATAATATGGTGCCAAATTTGATATCAGAAG 240
QY 241 TGGCTTCTAAACCAATGATATTTGCTGTGATGGGAGCAGTACTACTGCAACAGTTTGTGACAC 300
Db 241 TAGCTTCAAAACCAATGATATCGCAGGTGATGGAACCTACAACCTGCAACTGTTTGTGACC 300
QY 301 AAGCCATTGTTCTAGAGGACTTAAATAATGTGACAGCAGGTGCTTAATCCAATTCGTTATCC 360
Db 301 AGCAATCGTCGTTGAAGGAATCAAAACGCTCAGCAGGTGCAAAATCCAATCGGTATTC 360
QY 361 GTCGAGCATTTGAACACAGCAGCAGCTGTTGAAGCCTTTGAAAGCCATTTCTCTCAAC 420
Db 361 GTCGTTGGGATTTGAACACAGCAGTTGCCGACAGCTTGAAGCTTTGAAACACACGTCATCC 420
QY 421 CTGTATCTGGCAAGGAGCTATTGCTCAGGTGCTGCTGAGTATCATCATGCTCTCTGAAAAG 480
Db 421 CTGTTGCAATAAAGAGCTATCGTCAAGTTGACGCGCTATCTCTCTGTTCTCTGAAAAG 480
QY 481 TTGAGAGATATCTCAGAGCTATGAGCGCTGTCGCGCAACGATGCTGTGATTTACCATCG 540
Db 481 TTGTTGAGTACATCTCTGAAGCAATGGAAGTTGGCAAGCGGTGTCATCACCATCG 540
QY 541 AAGAATCTCGAGGTATGGAACACAGAACTTGAAGTGGTTGAAGGATCAATTTGACCGTG 600
Db 541 AAGAGTCACGTGATGGAACACAGAGCTTGAAGTGGTAGAAGGAATGCAGTTTGACCGTG 600
QY 601 GTTACCTGTCTAATATACATGCTCAGACAGATGAAATAATGGTTGCGAGCTTTGAAAACC 660
Db 601 GTTACCTTTCACAGTACATGCTGACAGATAGCGAAAAATGGTGGCTGACCTTGAAAATC 660
QY 661 CATTATCTTAATCAGCGATTAATAAGTGTCAACATCCCAAGCATTTTCCCACTACTTG 720
Db 661 CGTACATTTGATTACAGCAAGAAATTTCCAAATATCCCAAGAAATCTTCCCACTTTGG 720
QY 721 AGGAAGTCTTAAACCAACCGCTCATTACTTCAATTTGAGATGATGTTGATGGTGAAG 780
Db 721 AAGCATTTCTCAAGCAATCGTCCACTCTTGATTATTTGGGATGATGTTGATGGTGAAG 780
QY 781 CACTTCAACCCCTTGTCTTGAACAAAGATTCGTGTACTTTCAATTTGCTGTCTCAAAAG 840
Db 781 CTCTTCAACTCTGTTTGAACAAAGATTCGTGGAACCTTCAACGTTAGTAGCAGTCAAGG 840
QY 841 CGCCAGGATTTGTTGATCGTCTGTAAGCTATGCTTGAAGCATTTGCTATCTTGACAGGTG 900
Db 841 CACCTGTTTGTGTCGCGCAAGCCATGCTTGAAGATATGCGCATCTTAAACAGGCG 900
QY 901 GTACAGTATTACAGAGGATCTAGGACTTGAATTTAAAGATGCTACAAATGACAGCCCTTG 960
Db 901 GAACAGTTATCAGAGAAGACCTTGGTCTTGAGTTGAAGATGCGCAATTTGAAGCTCTTG 960
QY 961 GACAGGCTGCTAAGATTACAGTTGATTAAGATAGCAGCAATTTGTTGAAGTTTCAAGAA 1020
Db 961 GTCAAGCAGCGAGAGTGCACCGTGGACAAGATAGCAGCGTTATTGTTAGAAGTGCAGAA 1020
QY 1021 GTTCAGAGCATTTGCTAAACCGTATTCGCTGATTAATTCGCANTTAGAACAACACACTT 1080
Db 1021 ATCCTGAAGAGGATTTCTCACCGGTGTTGCGGTATTCAAAGTCTCAAAATGAAACTACA 1080
QY 1081 CTGACTTTGACCGTGAATAAATACTACAAGAGCTTTGGGCAAAATAGCTGGTGGTGTAGCTG 1140
Db 1081 CTGATTTGACCGTGAATAAATACTACAAGAGCTTTGGCAAAATGTCAGGTTGGTGTAGCGG 1140
QY 1141 TTATCAAAGTAGGAGCTTCAACACAGAGACAGCTTTTAAAGAAATGAAACTTTCGCAATGAG 1200
Db 1141 TTATTAAGGTGCGGAGCGCAACTGAAACTGAGTTGAAGAAATGAAACTCCGCAATTGAAG 1200
QY 1201 ATGCTCTAATGCTACAGTGCAGCGCTTGAAGAGGTATGCTGCTGCTGGTGGTGAACAG 1260
Db 1201 ATGCCCTCAACGCTTACTCGTGCAGCTGTTGAAGAGGTATTTGTTGCAGGTGGTGAACAG 1260

QY 1261 CACTTATACGGTATTGAAAAAGTAGCAGCTCTTGAGCTTGAGGCGGATGATGCTACTG 1320
 Db 1261 CTCTTGCCAAATGTGATTCAGCTGTTGCTACCTTGGAAATTTGACAGGAGATGAAGCAACAG 1320
 QY 1321 GAGGTACATTTGCTTCGTGCTCTAGAGAGGCTGTACGTCAAAATGCTTTAAATGCTTG 1380
 Db 1321 GAGGTAAATTTGCTTCGTGCTCTAGAGAGGCTGTACGTCAAAATGCTTTAAATGCTTG 1380
 QY 1381 GGTACGAAGGCTCCGTAGTATTGACAAGTTGAAAAACAGCCCTGCAGGAACAGGATTTA 1440
 Db 1381 GATTGGAAGATCTCGTTATCGATCGTTTGAATAATGCTGAGCTTGGTATAGGATTC 1440
 QY 1441 ATGCTGCAACAGAGTGTGGTGTATGATTAATAACAGGAATCATTTGACCCCTGTCAAAG 1500
 Db 1441 ACCGAGCAACTGCGAGTGGGTTAAATGATGATCAAGATGATCAATGATCCAGTTAAAG 1500
 QY 1501 TACACGATCAGGCTTCAAAATGAGCTTCTGATGATGATGATGATGATGATGATGATG 1560
 Db 1501 TGAGTGTGCTGAGCCCTACAAATGAGCTTCTGATGATGATGATGATGATGATGATGAT 1560
 QY 1561 CAGTGTGCTTAATAACCTGAAACCTGAGCTAGCCAGCCAGCAATGCCAGCAGGATGG 1620
 Db 1561 CAGTGTGCTTAATAACCTGAAACCTGAGCTAGCCAGCCAGCAATGCCAGCAGGATGG 1620
 QY 1621 ATCCAGGAATGATGG 1636
 Db 1621 TGGGTGGAATGGCGG 1636

RESULT 3
 US-09-583-110-1024
 ; Sequence 1024, Application US/09583110
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al.
 ; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
 ; FILE REFERENCE: PATH00-07A
 ; CURRENT APPLICATION NUMBER: US/09/583,110
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/107,433
 ; PRIOR FILING DATE: 1998-06-30
 ; PRIOR APPLICATION NUMBER: US 60/085,131
 ; PRIOR FILING DATE: 1998-05-12
 ; PRIOR APPLICATION NUMBER: US 60/051,553
 ; PRIOR FILING DATE: 1997-07-02
 ; NUMBER OF SEQ ID NOS: 5322
 ; SEQ ID NO 1024
 ; LENGTH: 1623
 ; TYPE: DNA
 ; ORGANISM: Streptococcus pneumoniae
 US-09-583-110-1024

Query Match 62.6%; Score 1039.2; DB 21; Length 1623;
 Best Local Similarity 77.6%; Pred. No. 8.7e-270;
 Matches 1257; Conservative 0; Mismatches 363; Indels 0; Gaps 0;

QY 15 ATGGCAAAAGAAATCAAAATTTTACAGAGATGCGGTGCTGCCATGGTGGCGGAGTGTAT 74
 Db 1 atgtcaaaagaaataaattttcatcagatgcccgctcagccatggttcgtggttcgat 60
 QY 75 ATGTTAGCAGATACCTCAAAATGAAAGTGGTTCCTTAAAGGGCGCAATGTTGTTCTTGA 134
 Db 61 attcttgacacactgttaagtaaccttggaccacaaaggtcgcaatgtcgttcttgaa 120
 QY 135 AAAGCTTTGGTTCCTCCCTTAATTAATGACGGGTAAACCATGCTTAAGAGATCGAA 194
 Db 121 aagtcattcgttcaccccttgattcccaatgacggtgtgaccatgccaagaagaatcgaa 180
 QY 195 TTAGAGATCATTTGAAACATGGACAAATTTGGTCTCTCAAGTGGCTTCTAAAC 254
 Db 181 ttggaagaccattttgaaaataatggtgctaaagtagtatcatcagaagtagcttctaaacc 240

QY 255 AATGATATTGCTGATGGAGCAGCTACTGCAACAGTTTGTGACAAAGCCATTGTTTCA 314
 Db 241 aatgatatcgaggtgacggaactacgactgcaagagtttgaccacaaagctatcgccgt 300
 QY 315 GAAGGACTAAAAAATGTGACAGAGTGTAAATCCAAATTTGGTATCGTTCGAGCAGCATGAA 374
 Db 301 gaaggaaatcaaaaactgacagcaggtgcaaatcccaatcggtatttcgtcggtggtgaa 360
 QY 375 ACAGCAACAGCAACAGCTGTTGAAGCTTTGAAGCCATTGCTCAACCTGTATCTGCAAG 434
 Db 361 acagcagttgcccagcagttgaagctttgaaaaaacagccatcccttggccaataaa 420
 QY 435 GAAGCTATTGCTCAGCTCGCTGAGTATCATCACGCTCTGAAAAAGTTGGAGAGTATATC 494
 Db 421 gaagctatcgctcaagtgcagcagtatctctcgttctgaaaaaagtgtgtgagtaac 480
 QY 495 TCAGAAAGCTATGAGGCGTGTGGCAACAGATGTTGATTAACCATCGAAGAAATCTCGAGGT 554
 Db 481 tctgaagcaatggaaaaagtggcaagacggtgtcatcaccatcgaaagagtcacgtggt 540
 QY 555 ATGGAACAGAACTTCAAGTGTGTTGAAGCATCCAAATTTGACCGTGGTTACCTGTCTCAA 614
 Db 541 atggaacacagacttgaaagtcgtagaagaaatgcagtttgaccgtggtcacctttcacag 600
 QY 615 TACATGCTCACAGCAATGAAAAATGTTGAGACCTTTGAAAAACCCATTATCTTAATC 674
 Db 601 tacatggtgacagatagcgaataatggtggtgacacttgaaataatccgtacatttgcatt 660
 QY 675 ACGGATAAAAAAGTGTCAAAACATCCAAAGACATTTTGCCACTACTTGAGGAAGTTCCTTAA 734
 Db 661 acagacaagaaaaattttcccaatcccaagaaatcttgccacttttggaagacatttcccaa 720
 QY 735 ACCAACCTGCCATTACTCATTTTTCAGATGATGTTGGATGGTTCAGACACTTCCAACCTT 794
 Db 721 agcaatcgctccactcttgatttcggatgatggtggtgagctcttccaaactctt 780
 QY 795 GTCTTGAAACAAGATCGTGTACTTTCAATGTGGTGTGCTCAAAAGCGCAGGATTTGTT 854
 Db 781 gttttgaaacaagatctgtgaaaccttcaacgtagtagcagtcgaagcaactggttttgg 840
 QY 855 GATCGTCTAAAGCTATGCTTGAAGACATTTGATGTTTGAAGAGTGGTACAGTGTATACA 914
 Db 841 gaccgtgcgaagccatgcttgaagataatcgcaatccttaacagcggaagacagttatcaca 900
 QY 915 GAGGATCTAGGACTTCAATTAAGATGCTACAATCAGACAGCCCTTGGACAGCTGCTAAG 974
 Db 901 gaagaccttggctcttgatgagtgcaaatgcaaatgcaaatgcaaatgcaaatgcaaatgca 960
 QY 975 ATTACAGTTGATAAAGATAGCAGTAAATTTGAAGGTTTCAGGAAGTTTCAGGAAGTATT 1034
 Db 961 gtgacctggaacaagatagcagcgttattgtagaagtgcaagaaatcctgaagcgtatt 1020
 QY 1035 GCTAACCGTATTGCTGCTGATTAATCGCAATTTAGAAACAACTTCTGACCTTTGACCGT 1094
 Db 1021 tctcacggtgtgctggttatcaagtctcaaatcgaaactacaaacttctgaatttgacct 1080
 QY 1095 GAAAACTACAAAGACGTTTGGCGAAATTAGCTGGTGGTGTAGCTGTTTATCAAGTAGGA 1154
 Db 1081 gaaaaatgcaagaacgcttggccaaattgtcaggtggtgtagcgttatttaaggtcgga 1140
 QY 1155 GCTCAACAGACAGCTTTTAAAGAAATGAACTTCGATTTGAGGATGCTCTTAATGCT 1214
 Db 1141 gccgcaactgaaactgagttgaaagaaatgaaactccgcattgaaagatgccctcaacgt 1200
 QY 1215 ACAGTGCAGCGGTTGAAGAAGGTATCGTGTGCTGGTGGGAACACACTTCTATTACGGTT 1274
 Db 1201 actcgtgcagctgttgaagaaggtattgtgaggtggtggaacagctcttgcgaatgtg 1260
 QY 1275 ATTGAAAAGTAGCAGCTTTCAGCTTGAAGCGGATGATGCTACTTGGAGCTAACATTGTTG 1334
 Db 1261 attcagctgtgttcaccttgaattgacagagatgaaacacagagcgttaattgtt 1320

QY 1515 CTTCAAAATGCAGCTTCTGTAGTACTTATTTTGAACACAGACAGTGTGTGCTAAT 1574
 Db 1501 ctacaaaatgcagcatctgtagccagctgattttgacacagagcagctgtagccaat 1560
 QY 1575 AAACCTGACACAGCTACGCCAGCGCCAGCAATGCCAGCAGGTATGATCCAGGAATGATG 1634
 Db 1561 aaaccagaaccagtagccccagctccagcaatggatcccaagtatgatggcgggagtagtg 1620

RESULT 5

US-09-613-303-50
 ; Sequence 50, Application US/09613303
 ; GENERAL INFORMATION:
 ; APPLICANT: Siegel, Marvin
 ; APPLICANT: Chu, N. Randall
 ; APPLICANT: Mizzen, Lee A.
 ; TITLE OF INVENTION: INDUCTION OF A TH1-LIKE RESPONSE IN VITRO
 ; FILE REFERENCE: 12071/002001
 ; CURRENT APPLICATION NUMBER: US/09/613,303
 ; CURRENT FILING DATE: 2000-07-10
 ; PRIOR APPLICATION NUMBER: US 60/143,757
 ; PRIOR FILING DATE: 1999-07-08
 ; NUMBER OF SEQ ID NOS: 55
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 50
 ; LENGTH: 1926
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: fusion sequence
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(1923)
 US-09-613-303-50

Query Match 62.5%; Score 1038.4; DB 22; Length 1926;
 Best Local Similarity 77.5%; Pred. No. 1.5e-269;
 Matches 1258; Conservative 0; Mismatches 366; Indels 0; Gaps 0;

QY 15 ATGGCAAAAGAAATCAAAATTTTACAGATGCGCTGCTGCCATGGTCCGGAGTGTAT 74
 Db 1 acggcaaaagaataataaatttcacatcagatgcccttcagctatgctgctggtgctgat 60
 QY 75 ATGTTAGCAGATACCGTCAAAGTAACGCTTGGTCTCTAAAGGCGCAATGTTCTTCTGAA 134
 Db 61 atcctgcagatcactgttaagtaactttgggacccaaggctcgcaatgctgtcttgaa 120
 QY 135 AAAGCTTTTGGTTCCTCCCTTAATTAATGACGGGGTAACCATTTGCTTAAGAGATCGAA 194
 Db 121 aagtcattcgttcccttgattcccaatgacggtgtgactatgccaagaagaattgaa 180
 QY 195 TTAGAGATCATTTTGAACATGGAGCAAAATTTGGTCTCTGAAGTGGCTTCTAAACC 254
 Db 181 ttagaagaccatttgaataatgggtgccaatttggatcagagagtagcttcaaaaacc 240
 QY 255 AATGATATTGCTGGTGGGAGCAGTACTGCAACAGATTTTTCACACAGCCATTGTTCAAT 314
 Db 241 aatgatatcgagtgatgtagaactcaactgcaactgttttgaccacgaacatgctcgt 300
 QY 315 GAAGACTTAAATAATGTGACACAGGTGCTAATCAATTTGGTATCCGTCGGAGGCATTTGAA 374
 Db 301 gaaggaaatcaaaacgtcacagcaggtgcaaatcccaatcgggtattcgtggtggattgaa 360
 QY 375 ACAGCAACAGCAACAGCTGCTTCAAGCCTTTGAAGCCATTGCTCAACCTGTATCTGGCAAG 434
 Db 361 acagcagttgcccgcagcagtgtaggctttgaaaaaacacacgtcatccctggtgccaatgaa 420
 QY 435 GAAGCTATTGCTCAGGTGCTGCTCAGTATCATCACGCTCTGAAAAAAGTTGGAGAGTATATC 494
 Db 421 gaagctatcgctcaagtgcagccgtatcttctctgcttctgtaaaagttgggtgagctacac 480
 QY 495 TCAGAAAGCTATGGAGCGGTGGGGCAACGATGGTGTGATTAACCATCGAAGATCTCGAGGT 554

Db 481 tctgaagcaatggaaaaagtgtggcaagacggtgtcatcaccatcgaaagtagtcaacgtggt 540
 QY 555 ATGGAAACAGAACTTCAACTGCTTGAAGCAATGCAATTTTACCGTGGTTACCTGTCTCAA 614
 Db 541 acggaacacagagcttgagctgtagaagaaagcagtttgacccgtggttacctttccacg 600
 QY 615 TACATGGTCCACAGCAATGAAAAATGTTGAGACGCTTGAACCCATTTATCTTTAATC 674
 Db 601 tacatggtgacagatagcgaaaaatgggtgctgaccttgaaaatccgtaattttgatt 660
 QY 675 ACGGATAAAAAGTGTCAACATCCAAACATTTTGGCCACTTCTTGAGGAAGTTCCTTAA 734
 Db 661 acagacagaagaaatttccaatccaagaatacttgcaccttttggaagacatcttccaa 720
 QY 735 ACCAACCGTCCATTACTTACTTATTCAGATGATGCGATGCGTGAACACCTTCCACCCCTT 794
 Db 721 agcaatcgccactcttgattatcggtgatgtggtggtgaggtcttccaactctt 780
 QY 795 GTCTTGACAAAGATTCGTGCTTCAATGTTGTTGCTGTCATAAGCGCCAGGATTTGCT 854
 Db 781 gtttgaacaagatctgtagaaccttcaacgttagtagcagtcgaagcactggttttgg 840
 QY 855 GATCGTGTAAAGCTATGCTTGAAGACATTTGCTATCTTGAACAGGTGTTACAGTATACA 914
 Db 841 gaccgtcgcaagccatgcttgaagatatcgccatctttaaagcgcggaacagttatcaca 900
 QY 915 GAGGATCTAGGACTTGAATTAAGATGCTACAATGACAGCCCTTTGGACAGCTGCTTAG 974
 Db 901 gaagaccttggcttggatgtagtgaagatcgcaaatggaagctcttggtaacagcagaga 960
 QY 975 ATTACAGTTGATAAGATAGACAGTAAATTTGTTGAAGTTTTCAGGAAGTTTCAAGATGAT 1034
 Db 961 gtgacctgggacaagatagacaggttattgtagaaggtgcaagaaatccttgaagcatt 1020
 QY 1035 GCTAACCGTATTGCTACTGATTAATTCGAATTAGAACAACAACTTCTGACTTTGACCGT 1094
 Db 1021 tctcacctggttgcggttatcaagtcctcaaatcgaaactacaacttctgaatttgacct 1080
 QY 1095 GAAAACTACAAAGAACGTTTGGCGAAATTAGCTGGTGGTGTAGCTGTTATCAAGTAGGA 1154
 Db 1081 gaaaaattgcaagaacgcttggccaaattgtcaggtggtgagcgttatttaaggtcgaga 1140
 QY 1155 GTCCAAACAGACAGCTTTTAAAGAAATGAACCTTCGATTGAGGATGCTCTAATGCT 1214
 Db 1141 gccgcaactgaaactgagttgaaagaaatgaaactccgcattgaaagatgccctcaacgt 1200
 QY 1215 ACAGCTGACCGCTTGAAGAGGTATGCTGCTGGTGGTGGGAACAGCATTATTTAGCGTT 1274
 Db 1201 actcgtcagctgttgaagaaggtattgttcaggtggtggaacagctcttgcgaatgtg 1260
 QY 1275 ATTGAAAAAGTAGCAGCTTCTGAGCTTGAGGCGGATGATGCTTGGAGCTAACATTTG 1334
 Db 1261 atccagctgttgcctaccttggaaatgacagagagatgagcagcaggtatattgtct 1320
 QY 1335 CTCTGCTCTAGAAGAGCTGTAGCTCAATTTGCTTTAAATGCTGGGTACGAAGGCTCC 1394
 Db 1321 ctcctgcttggagaagaccttgcgtcaaatgtctcaaatgcaagattgaaagatct 1380
 QY 1395 GTAGTTATTGACAAATTTGAAAAACAGCCCTTGCAAGCAACAGGATTTAATGCTGCAACAGT 1454
 Db 1381 atcgttatcgatcgtttgaaaaatgtagcttggtaggattcaacgcaagcaactggc 1440
 QY 1455 GAGTGGTGTGATGATTAATTAACAGGAATCATTTGACCTCTGTAAGATTAACAGTACAGG 1514
 Db 1441 gagtgggttaacatgattgatcaaggtatcatctgaccttcaaggttaagtgagctcagcc 1500
 QY 1515 CTTCAAAATGCAGCTTCTGTAGTACTTATTTTGAACACAGAAAGCAGTGTGCTTAAT 1574
 Db 1501 ctacaaaatgcagcatctgtagccagctttagccagctttagcaacagagcagctcgtagccaat 1560
 QY 1575 AAACCTGAACCACTACGCCAGCGCCAGCATGCGCAGAGGTATGATCCAGCAATGATG 1634

Db 1561 aaacagacacagtagcccgagctccagcaatgagatcccaagtagtggtggaatggc 1620
 QY 1635 GGTG 1638
 Db 1621 gga 1624
 RESULT 6
 US-08-961-527-77
 : Sequence 77, Application US/08961527
 : GENERAL INFORMATION:
 : APPLICANT: Charles Kunsch
 : TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
 : NUMBER OF SEQUENCES: 391
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Human Genome Sciences, Inc.
 : STREET: 9410 Key West Avenue
 : CITY: Rockville
 : STATE: Maryland
 : COUNTRY: USA
 : ZIP: 20850
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 : COMPUTER: HP Vectra 486/33
 : OPERATING SYSTEM: MSDOS version 6.2
 : SOFTWARE: ASCII Text
 : CURRENT APPLICATION DATA:
 : FILING DATE: US/08/961,527
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER:
 : CLASSIFICATION: 424
 : FILING DATE:
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Brookes, A. Anders
 : REGISTRATION NUMBER: 36,373
 : REFERENCE/DOCKET NUMBER: PB340P1
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (301) 309-8504
 : TELEFAX: (301) 309-8512
 : INFORMATION FOR SEQ ID NO: 77:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 5365 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: double
 : TOPOLOGY: linear
 : US-08-961-527-77

Query Match 62.3%; Score 1035.4; DB 13; Length 5365;
 Best Local Similarity 77.4%; Pred. No. 1.5e-268;
 Matches 1255; Conservative 0; Mismatches 366; Indels 0; Gaps 0;

QY 14 TATGGCAAAAGATCAATTTTACAGAGATGCGCGTGCCTGCCATGTCGCGCGAGTTGA 73
 Db 277 TATGTCAAAAGAAATTAATTTTATCATGATGCGCGTGCAGCCATGTTGCGGTGTCGA 336
 QY 74 TATGTAGCAGATACCGTCAAAAGTAAAGTTCCTTAAAGGCGCAATGTTGTTCTTGA 133
 Db 337 TATCCTTGCAGACACTGTTAAAGTAACTTTGGGACCAAAAGGTGCGAATGTCGTTCTGA 396
 QY 134 AAAAGCTTTTGGTTCCTCCCTTAATTAATGACGGGGTAACCATTTGCTAAAGAGATCGA 193
 Db 397 AAAAGTCATTCGGTTCACCCCTTGATTACCAATGACGGGTGTCACCATTTGCCAAAGAAATCGA 456
 QY 194 ATTAGAAGATCAATTTTCAAAACATGGGACGAAATTTGGTGTCTGGAAGTGGCTCTTAAAC 253
 Db 457 ATTGGGAAGACATTTTGAATAATGTTGGTGTCTAAGTTAGTATCAGAAGTAGCTTCTTAAAC 516
 QY 254 CAATGATATTGCTGGTGGGAGGACTACTGCAACAGTTTGTGACACAAAGCCATTGTTCA 313
 Db 517 CAATGATATCGAGGTGACGGAAGTACTGCAACAGTCTTGACCCAAAGCTATGTCGTCG 576

QY 314 TGAAGGACTAAAAATGTGACAGCAGGTGCTTAATCCAAATGGTATCGCTCGAGCATTTGA 373
 Db 577 TGAAGGAATCAAAACAGTCACAGCAGGTGCTTAATCCAAATGGTATCGCTCGAGCATTTGA 636
 QY 374 AACAGCAACAGCAGCTGTTGAAGCCTTGAAAGCATTGCTCAACCTGTFATCTGGCAA 433
 Db 537 AACAGCAGTTGCCCGCAGCAGTTGAAGCTTTGAAAAACAACGCCATCCCTGTTGCCAATAA 696
 QY 434 GGAAGCTATTGCTCAGGTGCTGAGTATCATCACGCTCTGAAAAAGTTGGAGAGTATAT 493
 Db 697 AGAAGCTATCGCTCAAGTTGCGAGCCGATCTCTTCGTTCTGAAAAAGTTGGTGGAGTAT 756
 QY 494 CTCAGAAGCTATGGAGCGTGTGGCAACGATGCTGATTTACCATCGAAGAAATCTCGAGG 553
 Db 757 CTCGAAGCAATGGAAAAAGTTGGCAAGAGCGGTGTCATCACCATCGAAGAGTACACGTGG 816
 QY 554 TATGGAACAGAACTTGAAGTGGTTGAAGGATGCAATTTGACCGTGTACCTGTCCTCA 613
 Db 817 TATGGAACAGAGCTTGAAGTGGTAGAAGGAAATGCGATTTGACCGTGTACCTTTTCA 876
 QY 614 ATACATGGTCACAGCAATGAAAAAATGGTTGCAGACCTTTGAAAAACCATTTATCTTAAT 673
 Db 877 GTACATGGTCACAGATAGCGAAAAAATGGTGGCTGACCTTGAAAAATCCGTACATTTTAT 936
 QY 674 CAGGATAAAAAAGTGTCAACATCCAAAGCAATTTGCCACTACTTTGAGGAAGTTCTTAA 733
 Db 937 TACAGACAAGAAAAATTTCCAAATATCCAAAGAAATCTTGCCACTTTTGGAAAGCATCTCCA 996
 QY 734 AACCAACCGTCCATTTACTATTATTCAGATGATGTCGATGGTGAAGCACTTCCAAACCT 793
 Db 997 AGCAATCGTCCACTCTTGATTTGCGGATGATGTGGTGGCGAGGCTCTTCCAACTCT 1056
 QY 794 TGTCTTGAACAAGATTCGTGTGCTTCAATGTGGTGTCTCAAAAGCGCCAGGATTTGG 853
 Db 1057 TGTCTTGAACAAGATTCGTGGAACCTTCAAGTAGTAGCAGTCAAGGCACTGTTGG 1116
 QY 854 TGATCGTCGTAAGCTATGCTTGAAGCAATGCTTATCTTGACAGGTGTACAGATGATAC 913
 Db 1117 TGACCGTCGCAAGCCATGCTTGAAGATATCGCCATCTTAACAGGCGCAAGTATATCAC 1176
 QY 914 AGAGGATCTAGGACTTGAATTAAGATGCTACAAATGACAGCCCTTGGCAGGCTGCTAA 973
 Db 1177 AGAAGACCTTGGTCTTGAGTTGAAGATGCGCAATTTCAAGCTCTTGGTCAAGCAGCGAG 1236
 QY 974 GATTACAGTTGATAAAGATAGCAGCAATTTGTAAGGTTTCAGAAAGTTTCAGAAAGTAT 1033
 Db 1237 AGTGACCGTGACAAAGATAGCAGCGTTATTTGTAAGAGTGCAGGAATCTTGAGCGAT 1296
 QY 1034 TGCTAACCGTATTGCACTGATTTAAATCGCAATTTAGAAACAACAACTTCTGACTTTGACCG 1093
 Db 1297 TTCTCACCGGTTTGGGTTATCAAGTCTCAAAATCGAAACTACAACTTCTGAAATTTGACCG 1356
 QY 1094 TGRAAACTACAGAAGCTTTGGCAATTTAGCTGGTGGTGTAGCTGTTATCAAGTAGG 1153
 Db 1357 TGAATAATTCGAAGACGCTTTGGCCAAATTTCTCAGGTGGTGTAGCGGTTATTAAGGTTGG 1416
 QY 1154 AGCTCCACAGAGACAGCTTTTAAAGAAATGAAACTTCGCAATTTGAGGATGCTCTAAATGC 1213
 Db 1417 AGCGCAACTGAACTGAGTTGAAAGAAATGAAACTTCGCAATTTGAGGATGCTCTAAATGC 1476
 QY 1214 TACAGTGCAGCGCTTGAAGAGGATATCGTTGCTGGTGGGAAACAGCACTTATTACGGT 1273
 Db 1477 TACTCGTGCAGCTTTGAAGAAAGGTTATTTGTCAGGTGGTGGAAACAGCTCTTGGCAATGT 1536
 QY 1274 TATTGAAAAAGTACGACCTTTGAGCTTTGAGGGGATGATGCTACTGGACGCTACATTTGT 1333
 Db 1537 GATTCCAGCTGTTGCTACCTTGGAAATTTGACAGGAGATGAAAGCAACAGGACGTAATTTGT 1596
 QY 1334 GCTTCGTGCTTAGAAGAGCTGTACGTCAAATTTGCTTTTAAATGCTGGTACGAGGCTC 1393
 Db 1597 TCTCCGTGCTTTGGAAGAACCCGTTGCTCAAAATGCTCAACATGCAAGATTTGAAGATC 1656
 QY 1394 CGTAGTTATTTCACAAAGTTGAAAAACAGCCCTTGCAGGAACAGGATTTAATGCTCAACAGG 1453


```

Db 1657 TATCGTTATCGATCGTTGAAATAATGCTGAGCTGTGTATAGGATTTAAACGACGAACTGG 1716
QY 1454 TGAATGGGTTGATATCAATTAACAGGAATCATTTGACCCCTGTCAAAAGTAACACGATCAGC 1513
Db 1717 CGAGTGGGTACATGATGATCAAGTATCATTTGATCCAGTTAAAGTGAAGTGGTTCAGC 1776
QY 1514 GCTCAAAATGAGCTTCGTAGCTAGTCTTATTTTGAACAAGAGCAGTGTGCTAA 1573
Db 1777 CCTACAAATGAGCATCTGTAGCCAGCTTGATTTTGAACAAGAGCAGTGTAGCCAA 1836
QY 1574 TAACTGACAGCTACCCAGCCAGCCAGCAATGCCAGCAGTATGATCCAGGAATGAT 1633
Db 1837 TAAACAGAACAGTAGTACGCCAGCTCCAGCAATGGATCCAGCATGATGGCGGGATGAT 1896
QY 1634 G 1634
Db 1897 G 1897

RESULT 7
US-60-061-998-525
; Sequence 525, Application US/60061998
; GENERAL INFORMATION:
; APPLICANT: LAGACE, ROBERT E.
; APPLICANT: CORLEY, NEIL C.
; APPLICANT: RUSSO, FRANK D.
; APPLICANT: HANN, AMY L.
; APPLICANT: HEATH, JOE D.
; APPLICANT: FINNEY, GREGORY L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF STREPTOCOCCUS PNEUMONIAE
; TITLE OF INVENTION: GENOME, FRAGMENTS THEREOF, AND USES THEREOF
; NUMBER OF SEQUENCES: 797
; CORRESPONDENCE ADDRESS:
; ADDRESS: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/061,998
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PM-0006-2P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 525:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3714 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: SPN1C530
US-60-061-998-525

Query Match 61.7%; Score 1024.4; DB 31; Length 3714;
Best Local Similarity 77.4%; Pred. No. 1.2e-265;
Matches 1255; Conservative 0; Mismatches 366; Indels 1; Gaps 1;
QY 14 TATGCCAAGAAATCAAAATTTTACAGCATGCGCGTGTGCCATGGTGGCGGAGTTGA 73

```

```

Db 224 TATGTCAAAAGAAATTAATTTTCATCAGATGCCCGTTCAGCCATGTTCTCGTGGTTCGA 283
QY 74 TATGTTAGCAGATACCGTCAAAAGTAACGCTTGCTCCCTAAAGGGCGCAATGTTCTTCTTGA 133
Db 284 TATCCTTGCAGACACTGTTAAAGTAACCTTGGAGCCAAAGGTGCGCAATGTCGTTCTTGA 343
QY 134 AAAAGCTTTTGGTTCCTCCCTTAATTAATGACGGGGTAACCATTTGCTAAAGAGATCGA 193
Db 344 AAAGTCATTTCGGTTTCAACCTTGATTACCAATGACGCTGTGACCATTTGCCAAGAAATCGA 403
QY 194 ATTAGAAGATCATTTTGAACACATGAGGAGCAAAATTTGGTGTCTGAAAGTGGCTTCTTAAAC 253
Db 404 ATTGGAAGACCATTTTGAATAATATGGTGTCTAAGTTAGTATCAGAAGTAGCTTCTTAAAC 463
QY 254 CAATGATATTGCTGGTGTGGGACGAC-TACTGCAACACATTTTGCACACAGCATTTGTTTC 312
Db 464 CAATGATATCGCAGGTGACGGGACTACGGACTGCAACAGCTTTGACCCCAAGCTATCGTCC 523
QY 313 ATGAGGACTAAAAATGTGACAGCAGGTGCTAATCCAATTTGGTATCCGTCGAGGCATTG 372
Db 524 GTGAAGGAATCAAAACGTCACAGCAGGTGCAATCCAATCGGTATTCGTCTGGTGGATTG 593
QY 373 AAACAGCACAGACAGCTGTTGAAGCCTTGAAGCCATTTGCTCAACCTGTATCTGGA 432
Db 584 AAACAGCAGTTCGCCGACGAGTTGAAGCTTTGAAAAACAACGCCATCCCTGTTGCCAATA 643
QY 433 AGSAGCTATTGCTCAGGTCGCTGACGATATCATCAGCTCTGAAAAAGTTGGAGAGTATA 492
Db 644 AAGAAGCTATCGCTCAAGTTGACGCGGTATCTCTCGTCTGAAAAAGTTGGTGGAGTACA 703
QY 493 TCTCAAGACTATGAGCAGCTGTGGGCAACGATGGTGTGATTACCATCGAAGATCTCGAG 552
Db 704 TCTCTGAAGCAATGGAAAAAGTTGGCAAGACGGTGTCTATCACCATCGAAGAGTCAAGTG 763
QY 553 GTATGGAACAGAACTTGAAGTGGTTGAAGGCATGCAATTTGACCGTGGTTACCTCTCTC 612
Db 764 GTATGGAACAGAGCTTGAAGTGGTGAAGAGGAATGCAAGTTGACCGTGGTTACCTCTTCA 823
QY 613 AATACATGTCACAGACAATGAAAAAATGTTGACAGACCTTGAAAAACCCATTTATCTTAA 672
Db 824 AGTACATGGTGACTGATACGGAATAATGGTGGTGGCTGACCTTGAAAAATCCGTACATTTGA 893
QY 673 TCACGGATAAAAAAGTGTCAAAACATCCAAAGACATTTTGGCAGCTACTTGAAGAGTTCTTA 732
Db 884 TTACAGACAAGAAATTTCCATATCCAAAGAAATCTTGCCACTTTTGGAAAGCATTTCTCC 943
QY 733 AAACCAACCGTCCATTACTCATTTATTTGACAGATGATGATGGTGAAGCAGCTTCCAACCC 792
Db 944 AAAGCAATCGTCCACTCTTGATTATTTGCGGATGATGATGGATGCGGAGCTCTTCCAACATC 1003
QY 793 TTGCTCTTGAACAGATTCGTGCTACTTTCAATGTTGGTTCCTCAAGCGCCAGGATTTG 852
Db 1004 TTGTTTGAACAGATTCGTGGAACCTTCAACGATAGTAGCAGTCAAGGACCTCTGTTGTTG 1063
QY 853 GTGATCGTCTAAAGCTATGCTTTGAAGACATTTGCTATCTTTGACAGGTGGTACAGTATTA 912
Db 1064 GTGACCGTCGAAAGCCATGCTTTGAAGATATCGCCATCTTACAGCGGGAACATATCA 1123
QY 913 CAGAGATCTAGGACTTGAATTTAAAGATGCTPACATGACAGCCCTTGGACAGGCTGCTA 972
Db 1124 CAGAAGACCTTGGTCTTGAGTTGAAAGATGCGACAATTTGAAGCTCTTGGTCAAGCAGCA 1183
QY 973 AGATTACAGTTTGAAGATGACAGATTAATTTGTAAGGTTTCAAGAGTTTCAAGAGCTA 1032
Db 1184 GAGTACCGTGGACAAGATAGCAGGTTATTGTAAGAGGTGCGAAGAAATCTCGTAAGCGA 1243
QY 1033 TTGCTAACCGTATTGCAGCTGATTAATGCAATTTAGAAACAACACTTCTGACTTTGACC 1092
Db 1244 TTTCTCACCGTCTGGGTTATCAAGTCTCAATCGAAGAACTACAACTTCTGAATTTGACC 1303
QY 1093 GTGAAAAACTTACAAGACGTTTGGCGAAATTAGCTGGTGGTGTAGTGTATTATCAAGTAG 1152

```



```
Db 1304 GTCAAAATTTGCAAGACGCTTGGCCAAATTTGCAGTGTGTAGCGGTTATTAAAGTTG 1363
QY 1153 GAGCTCCACAGACAGACGCTTTAAAGAAATGAATCTCGATTGAGGATGCTCTAAATG 1212
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1364 GAGCCCAACTGAACTGAGTTGAAAGAAATGAATCTCGCATTTGAAGATGCCCTCAACG 1423
QY 1213 CTACAGCTGCAGCGCTTGAAGAGGTTATCGTTCTGTTGGTGAACAGCACTTATTACGG 1272
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1424 CTACTCTGCGAGCTGTTGAAGAAAGTATTGTTGACAGTGTGTGAACAGCTCTTTGCCAATG 1483
QY 1273 TTATTGAAATGAGCAGCTCTTGAAGCTTGAGGCGGATGATGCTACTTGGACGTAACATTG 1332
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1484 TGATTCAGCTGTGTACCTTGGAAATTGACAGAGATGAAGCAACAGACGTCGTAATATTG 1543
QY 1333 TGCTTCGTGCTCTAGAAGACGCTGTACGTCAAAATTTGCTTTAAATGCTGGGTACGAAGGCT 1392
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1544 TTCTCGTGCTTTGGAAGAACCGCTTCGTCAAAATGCTCAAAATGCTCACAATTTGAAGGAT 1603
QY 1393 CCGTAGTTATTGACAAGTTGAAAGAACAGCCCTCCAGGAACAGGATTTTAAATGCTGCAACAG 1452
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1604 CTATCGTTATCGATCGTTTGAAGAAATGCTGAGCTTGGTATAGGATTTAAGCGCAACTG 1663
QY 1453 GTGAGTGGGTTGATGATTAAAGACAGGAATCATTTGACCCCTGTCAAGATGAACAGATCAG 1512
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1664 GCAGTGGGTTAAATGATGATCAAGGATCATTTGATCCAGTTAAAGTGAAGTGGTTTCAG 1723
QY 1513 CGCTTCAAAATGAGCTCTCTGAGTGTGTTTAAAGCAAGAGCAAGTGTGTTGCTA 1572
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1724 CCTCAAAATGAGCATCTGTAGCCAGCTTGATTTTGACACAGAGCAGTGTAGCCA 1783
QY 1573 ATAACTGAAACAGCTACGCCAGCCAGCCAGCAATGCGCAGGATGATGATCCAGGAATGA 1632
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1784 ATAAAGCAGAACAGTAGAGCCAGCTCCAGCAATGGATCCAAAGCATGATGGCGGGATGA 1843
QY 1633 TG 1634
Db 1844 TG 1845
```

RESULT 8

```
US-60-029-960-223
; Sequence 223, Application US/60029960
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 1649
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/029,960
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 223:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5373 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-60-029-960-223

Query Match          61.6%; Score 1023.4; DB 27; Length 5373;
Best Local Similarity 77.4%; Pred. No. 2,7e-265;
Matches 1254; Conservative 0; Mismatches 366; Indels 1; Gaps 1;

QY 14 TATGGCAAAAGAAATCAAAATTTTCAGCAGATGCGCGTGTGCCATGTGCGCGGAGTTGA 73
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 272 TATGTCAAAAGAAATTAATTTTCATCAGATGCGCGTGTCCAGCCATGTTGTTGGTGTGA 331
QY 74 TATGTTAGCAGATACCGTCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 133
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 332 TATCCTTGGCAGACACTGTTAAAGTAACTTTGGGACCAAAAGGTCGCAATGTCTTCTTGA 391
QY 134 AAAAGCTTTTGGTCTCCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 193
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 392 AAAGTCATTCGGTTCACCTTGATTACCAATGAGCGGTGTGCCATTTGCCAAGAATCGA 451
QY 194 ATTAGAAGATCAATTTGAAACATGGGAGCAAAATTTGGTGTCTGAAAGTGGCTTCTTAAAC 253
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 452 ATTGGAAGACCAATTTGAAATATATGGTGTAAAGTATAGTATCAGAAGTAGCTTCTAAAC 511
QY 254 CAATGATATTTGCTGGTGTATGGGAGGACTACTGCAACAGTTTTTGACACAAGCCATTGTCA 313
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 512 CAATGATATTCGAGGTGACGGAACCTAGACTGCAACAGTCTTGACCCAAAGCTATTCGTCG 571
QY 314 TGAAGGACTAAAAAATGTGACAGCAGGTGCTAATCCAAATTTGGTATCCGTGAGGCAATTGA 373
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 572 TGAAGNATCAAAACGTCACAGCAGGTGCAATCCAAATCGGTATTCGTCTGGGATTTGA 631
QY 374 AACAGCAACCAACAGCTCTTGAAGCCTTGAAGCCTTGAAGCCTTGAAGCCTTGAAGCCTT 433
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 632 AACAGCAGTTGCGCGCAGCAGTTGAAGCTTTGAAAAAACAAGCCCTCCCTGTGCAATAA 691
QY 434 GGAAGCTATTGCTCAGGTCGCTGCAGTATCATCAGCTCTGAAAAAGTTGGAGAGTATAT 493
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 692 AGAAGCTATCGCTCAAGTTGCGAGCCGTTCTCTCGCTCTGAAAAAGTTGGTGAAGTACAT 751
QY 494 CTCAGAGCTATGAGCGGTGTGGGCAACGATGGTGTGATTACCATCGAAGAAATCTCGAGG 553
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 752 CTCTGAGCAATGGAANAAGTTGGCAAGAGCGGTGTCTCACCATCGAAGAGTCACTGG 811
QY 554 TATGGAACAGAACTTGAAGTGGTTGAAGCATGCAATTTGACCGGTGTTACCTGTCTCA 613
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 812 TATGGAACAGAGCTTGAAGTCTGAGTAGAAGGATGCAAGTTTGAACGTTTACCTTTTCA 871
QY 614 ATACATGGTCAACAGCAATGAAAAAATGGTTGCAGACCTTTGAAAAACCCATTATCTTAAT 673
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 872 GTACATGGTGCAGACATAGCGAAAAAATGGTGGCTGACCTTTGAAAAATCCGACATTTTAT 931
QY 674 CACGGATAAAAAAGTGTCAAAACATCCAAAGACATTTTCCCACTACTTTCAGGAAGTCTTAA 733
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 932 TACAGACAAGAAATTTCCAAATTCCAAGAAATCTTGCCACTTTTGGAAAGCATCTTCCA 991
QY 734 AACCAACCGTCCATTACTCAATTTGCAAGATGATGTGGATGGTGAAGCACTTCCAAACCT 793
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 992 AAGCAATCGTCCACTCTTGATTTATGGGATGATGTGGATGGGAGGCTCTTCCAACTCT 1051
QY 794 TGCTTTGAACAAGATTCGTGGTACTTTCAATGTGGTGTCTGTCACAGCGCAGGATTTGG 853
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1052 TGTTTGAACAAGATTCGTGGAAACCTTTCAACGTA-TAGCAGTCAAGGCACTGTTTGG 1110
QY 854 TGATCGTGTAAAGCATGCTTGAAGACATTTGATCTTGTGACAGGTGTACAGTGTATAC 913
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1111 TGACCGTCGCAAGCCATGCTTGAAGATATCGCCATCTTAAAGGCGGACAGTTATCAC 1170
QY 914 AGAGGATCTAGGACTTTGAATTTAAAGATGTACAATGACAGCCCTTGGACAGGCTGCTAA 973
```

```

;
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/068,175
; FILING DATE: HERewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PM-0009-2 P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-416
; INFORMATION FOR SEQ ID NO: 420:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2597 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: genomic DNA
; IMMEDIATE SOURCE:
; CLONE: SPN2c4.22
;
US-60-068-175-420

```

```

;
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/068,175
; FILING DATE: HERewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PM-0009-2 P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-416
; INFORMATION FOR SEQ ID NO: 420:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2597 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: genomic DNA
; IMMEDIATE SOURCE:
; CLONE: SPN2c4.22
;
US-60-068-175-420

```

```

Query Match 61.2% Score 1016.6; DB 31; Length 2597;
Best Local Similarity 77.4%; Pred. No. 1.4e-263;
Matches 1257; Conservative 0; Mismatches 364; Indels 2; Gaps 2;

QY 14 TATGCGAAGGAAATCAAAATTTTCAGCAGATGCGGCTGCTGCGATGGTGGCGGGAGTTGA 73
Db 230 TATGTCAGGAAATTAATTTTCATCAGATGCCGCTTCAGCTATGCTCGGTGCTCCA 289
QY 74 TATGTTAGCAGATCCGCTCAAAAGTAAAGCTTGGTCTTAAAGGGCGCAATGTTGTTCTGA 133
Db 290 TATCTTGCAGATCTGTTAAAGTAAAGCTTGGGACCAAAAGGTCGAATGCTGTTCTGA 349
QY 134 AAAAGCTTTGGTCTCCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 193
Db 350 AAATCAATTTGGCTCACCATTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 409
QY 194 ATTAGAAGATCATTTTGAAGACATGGAGCAAAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 253
Db 410 ATTAGAAGACCATTTTGAAGATATGGTGCCAAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 469
QY 254 CAATCATATTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 312
Db 470 CAATCATATTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 529
QY 313 ATGAAGGACTAAAAATTTGACAGCAGGTTGCTAAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 372
Db 530 GTGAAGGAAATCAAAACCGTCACAGCAGGTTGCAAAATTCRAATCGGTATTGCTGTTGGGATTG 589
QY 373 AAACAGCAACAGCAGCTGTTGAAGCTTTGAAGCCCTTGAAGCCATTTGCTCAACCTGTTGTTGTTGTT 432
Db 590 AAACAGCAGTTGCGCGCAGCAGTTGAAGCTTTGAAACCAACCAACCTGTTGTTGTTGTTGTTGTTGTT 649
QY 433 AGGAAGCTATTGCTCAGGTCGCTGAGTATCATCAGCTCTGAAAAGTTGAGAGTATA 492
Db 650 AAGAAGCTATCGCTCAAGTTGACGCGGTTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 709
QY 493 TCTCAGCAAGCTATGAGCGCTGTTGGCAACGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 552
Db 710 TCTCAGCAAGCTATGAGCAAGCTGTTGCAAGAGCGTGTATCATCAGCTGAGAGTACAGTG 769
QY 553 GTATGAAACAGCACTTGAAGTGGTTGAAGGCAATTTGACCGTGGTTACCTGTTGTTGTTGTTGTTGTTGTT 612
Db 770 GTATGAAACAGCAGCTTGAAGTGGTTGAAGGCAATTTGACCGTGGTTACCTGTTGTTGTTGTTGTTGTTGTT 829
QY 613 AATACATGTTACAGACAAATGAAAAATTTGTTGAGACCTTGAAGACCTTGAAGACCTTGAAGACCTTGAAG 672

```

```

RESULT 9
US-60-068-175-420
; Sequence 420, Application US/60068175
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert E.
; APPLICANT: Corley, Neil C.
; APPLICANT: Russo, Frank D.
; APPLICANT: Hann, Amy L.
; APPLICANT: Heath, Joe D.
; APPLICANT: Finney, Gregory L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF STREPTOCOCCUS PNEUMONIAE
; TITLE OF INVENTION: GENOME, FRAGMENTS THEREOF, AND USES THEREOF
; NUMBER OF SEQUENCES: 1175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:

```

```
Db 830 AGTACATGTTGACAGATAGCGAAAAAATGTTGGCTGACCTTGAATAATCCGTACATTTGA 889
QY 673 TCACGGATAAAAAAGTGTCAACATCCAGACATTTTGGCCACTACTTGGAGAAATGTTCTTA 732
Db 890 TTACAGACAAGAAAATTTCCAATATCCAAGAGATCTTGCCACTTTTGGAAAGCATCTCC 949
QY 733 AAACCAACCGTCCATTTACTCATTTATTCAGATGATGGATGGTGAAGCACTTCCAAACC 792
Db 950 AAAGCAATCGTCCACTCTTGATTTATTCGGATGATGGATGGCGAGGCTCTTCCAACTC 1009
QY 793 TTGCTTTGAACAAGATTCTGGTGTACTTTCAATGTGGTTG-CTGTCAAAGCGCCAGGATTT 851
Db 1010 TTGCTTTGAACAAGATCTGGTCAACTTCAAGCTAGTAGTCAAGCAAGCACTGGTTTT 1069
QY 852 GGTGATCGTGTGAAGCTATGCTGTGAAGACATTCGCTATCTTGACAGGTGGTACAGTGATT 911
Db 1070 GGTGACCGTGCAGAACCATCTGTTGAAGATATCGCCATCTTAACAGCGGAACAGTTATC 1129
QY 912 ACAGAGATCTAGGACTTGAATTAAGAGATGCTACATGACAGCCCTTGGACAGGCTGCT 971
Db 1130 ACAGAGACCTTGGTCTTGAGTTGAAGATGCGACAATTAAGAGCTCTGTGTCAGCAGCG 1189
QY 972 AAGATTACAGTTGATAAAGATACACAGTAATTTGTTGAAGTTTCAGGAAGCTTCAGAAAGCT 1031
Db 1190 AGAGTACCGTGCACAAAGATACACCGGTTATTTGTGAAGCTCCAGGAATCTGAAAGCG 1249
QY 1032 ATTGCTAACCGTATTGCACTGATTAAATCGCAATTAAGAAACAACAACCTCTGACTTTGAC 1091
Db 1250 ATTTCTACCGCTGTGCGGTTATCAAGTCTCAAAATCGAACTACAACCTCTGAATTTGAC 1309
QY 1092 CGTGAACAACTACAAGACGTTTGGCGAAATAGCTGGTGGTGTAGCTGTATCAAAAGTA 1151
Db 1310 CGTGAACAAATGCAAGAACGTTTGGCAAAATTTGTCAAGTGTGAGCGGTTATTAAGGTC 1369
QY 1152 GGAGCTCCACAGACAGCTTTAAAGAAATGAATTCGCAATTTAGGATGCTCTAAAT 1211
Db 1370 GGAGCCCAACTGAACTGAGTTGAAGAAATGAATTCGCAATTTAGGATGCTCTAAAT 1429
QY 1212 GCTACAGCTGACGCGGTTGAAGAGTATCGTTGCTGGTGGTGAACAGCACTTATTACG 1271
Db 1430 GCTACTCGTGCAGCTGTTGAAGAGGTTATTTGTCAGCGGTTGAACAGCTCTTGCAAT 1489
QY 1272 GTTATTGAAGAAAGTACGAGCTCTGAGCTTGAGGGGATGATGCTACTGACGCTAACAT 1331
Db 1490 GTGATTCAGCTGTGCTACCTTTGGAATTTGACAGGAGATGAAGCAACAGGACGTAATTT 1549
QY 1332 GTGCTTCGCTCTAGAGAGCCTGTACGTCAAAATTTGCTTTAAATGCTGGGTACGAAGGC 1391
Db 1550 GTTCTCGGCTTTGGAGAACCCGTTGTCAAATTTGCTCACAATGACGAGGATTTGAAGGA 1609
QY 1392 TCCGTAGTTATTGACAAGTTGAAAAACAGCCCTGCAGGAACAGGATTTAATGCTGCAACA 1451
Db 1610 TCTATCGTTATCGATCGTTTGAATAATGCTGAGCTTGATAGGATTTAAGCGAGCAACT 1669
QY 1452 GGTGAGTGGGTTGATATGATTTAAACAGGAATCATTTGACCCCTGTCAAAGTAACAGCATCA 1511
Db 1670 GGGAGTGGGTTAAACATGATTGATCAAGGTATCATTTGATCCAGTTAAAGTGAAGTCA 1729
QY 1512 GCGCTTCAAAATCGAGCTCTGTAGCTAGTCTTTATTTTGAACAACAGAGCAGTTGTTGCT 1571
Db 1730 GCGCTTCAAAATCGAGCTCTGTAGCCAGCTTGTATTTTGAACAACAGAGCAGCTGTAGCC 1789
QY 1572 AATAAACCTGAACAGCTAGCCGAGCCAGCAATGCGCAGGATGATGATCCAGGAATG 1631
Db 1790 AATAAACCAAGACAGTAGCCAGCTTCCAGCAATGATGATCCAGTATGATGGGGGATG 1849
QY 1632 ATG 1634
Db 1850 ATG 1852
```

RESULT 10
US-60-068-217-910

```
; Sequence 910, Application US/60068217
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert E.
; APPLICANT: Corley, Neil C.
; APPLICANT: Russo, Frank D.
; APPLICANT: Hann, Amy L.
; APPLICANT: Heath, Joe D.
; APPLICANT: Finney, Gregory L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF ENTEROCOCCUS FAECALIS
; NUCLEOTIDE SEQUENCES OF ENTEROCOCCUS FAECALIS
; GENOME, FRAGMENTS THEREOF, AND USES THEREOF
; CORRESPONDENCE ADDRESS: 1239
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/068,217
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PM-0003-5 P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 910:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8787 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: genomic DNA
; IMMEDIATE SOURCE:
; CLONE: EFALC910
; US-60-068-217-910

Query Match 56.9%; Score 945.8; DB 31; Length 8787;
Best Local Similarity 74.1%; Pred. No. 3.2e-244;
Matches 1196; Conservative 0; Mismatches 417; Indels 0; Gaps 0;

QY 15 ATGCGAAAGAAATCAAAATTTTCAGCAGATCGCGTCTGCTGCATGGTCGCGGAGTTGAT 74
Db 403 ATGCGAAAGAGATTAATTTGCAGAGATCGAGTCGAGCAATGCTACGCGGAGTAGAT 462
QY 75 ATGTTAGCAGATCCGTCAAAGTAACCGTTGGTCTTAAAGGGCGCAATGTTGTTCTTGA 134
Db 463 GTATTAGCAGATACAGTGAAGTGACATTAGGCGCTTAAAGTCTGTAACGTTGTTTGA 522
QY 135 AAAGCTTTTGGTCTCCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 194
Db 523 AAATCATTTTGGTTCACCATTTGATTACTTAACGATGGGGTAAACAATTTGTAAGGAA 582
QY 195 TTAGAAGATCATTTTGAACATGGGAGCAAAATTTGCTGTGAAGTGGCTTCTTAAAC 254
Db 583 TTAGAAGATCATTTTGAACATGGGAGCAAAATTTGCTGTGAAGTGGCTTCTTAAAC 642
QY 255 AATGATATTGCTGGTATGGGAGCACTACTGCAACAGTTTGTGACACAGCAATTTGTT 314
Db 643 AATGATATTGCTGGTATGGGAGCACTACTGCAACAGCAATTTGTTGACACAGCAAT 702
QY 315 GAAGGACTAAAAATGTGACAGAGGCTGCTAATCCATTTGTTATCCGTCGAGGCAAT 374
Db 703 GAAGGCTTAAAAACGTAACGCTGCTGAGCTTAACCCATTAGGTATTCGCCGTTGGA 762
```


Db	523	AAATCATTTGGTTCCACCATTGATTACATACGATGGGGTAAACAATTGCTAAGGAAATTTGAA	582
Qy	195	TTAGAAGATCATTTTGAAAACATGGGAGCAAAATTTGGTCTGAAGTGGCTTCTTAAACCC	254
Db	583	TTAGAAGATCATTTTGAAAACATGGGCGCAAAATTTAGTTTTCAGAAGTTGCTTCTTAAACG	642
Qy	255	AAATGATATTGCTGGTGATCGGGACGACTACTGCAACAGTTTTCACACAAGCCATTGTTTCAT	314
Db	643	AAATGATATTGCTGGTGACGGGAACAACACAGCAGCTGTTTTTGACACAAGCCATTGTTTCGT	702
Qy	315	GAAGGACTAAAAAATGTGACACAGGTGCTAATCCAAATTTGGTATTCGCTCGAGGCATTGAA	374
Db	703	GAAGGCTTAAAAAACGTAAGTCTGCTGAGCTAACCCATTAGGTATTTGCCGCTGGGATTTGAA	762
Qy	375	ACAGCAACAGCAACAGCTGTTGAAGCCTTGAAGCCATTGCTCAACCTGTATCTGGCAAG	434
Db	763	TTAGCAACAAAAACACAGCTAGAAGAAATTACACAATATTTTCATCTGTAGTTGATTTCAAAA	822
Qy	435	GAAGCTATTGCTCAGTCCGCTCGAGTATCATACGCTCTGAAAAGTTGGAGAGTATATTC	494
Db	823	GAAGCGATTGCAACAAGTCGCTGCTTTCATCAGGTTCTGAAAAGTTCGGCAAATTAATTT	882
Qy	495	TCAGAAGCTATTGAGCGTGTGGGCAACGATGCTGTCATTCACATCGAAGATCTCGAGGT	554
Db	883	GCCGATGCAATGAAAAGTTGGTACGACGCGGTAATTTAGCAATTTAGGAATCAAAAGG	942
Qy	555	ATGGAACACAGAACTTCAAGTGGTTGAAGCAATGCAATTTGACCGTGGTTACCTGCTCTCAA	614
Db	943	ATTGAAACAGAAATTAGATCGGTTGAGGAATGCAATTTGACCGCGGTTATTTATCTCAA	1002
Qy	615	TACATGGTCAAGACAATGAAAATAATGGTTGACACCTTTGAAAACCCATTATCTTTAATC	674
Db	1003	TACATGGTTACTGACAACGATAAATGGAAGCTGTTTGTAGAAAATCCATATATCTTTAATTT	1062
Qy	675	ACGGATAAAAAAGTGTCAAACTCCAAGACATTTTGCACACTACTTTGAGGAGTTCTTTAAA	734
Db	1063	ACCGACAAAAAATCTCAAATATTCAGGATATCTTACCTTTATTAGAACAAATTTCTACAA	1122
Qy	735	ACCAACCGTCCATTACTACTATTTGACAGATGATGGATGGTGAAGACATTTCAACACCTTT	794
Db	1123	CAAAGCGTCCACTATTTGATTATTCGGGATGATGTTGATGGGAAGCTCTACCAACATTTA	1182
Qy	795	GTCTTGAAACAAGATTGCTGGTACTTTCAATGTGGTTGCTGTCAAGGCGCCAGGATTTGGT	854
Db	1183	GTATTGAAACAAAATCCGTGGTACATTTAATGTTGTCGAGTAAAAGCGCCAGGATTTGGT	1242
Qy	855	GATCGTCGTAACACTATGCTTGAAGACATTTGCTATCTTGACAGGTGGTACAGTATTTACA	914
Db	1243	GACCGCGCAAGCGATGCTTGAAGATATGCTATTTTAAACAGGTGGTACAGTAATCACT	1302
Qy	915	GAGGATCTAGGACTTCAAATTAAGAATGCTACAATGACAGCCCTTGGACAGGCTGCTTAAG	974
Db	1303	GACGCTTAGGGTTAGAGTTTAAAGACACACTATTTGAAAATTTAGGAATTTGCTAGCCAA	1362
Qy	975	ATTACAGTTGATAAAGATAGCACAGCTAAATGTTGAAGTTTTCAGGAAGTTTCAGAAAGCTATT	1034
Db	1363	GTAGTTGCTGCACAAACATAACACAATTTGTCGAAGGTGCTGGTTCAAAGAAGCCATT	1422
Qy	1035	GCTAACCGTATTGCACTGATTTAAATGCGAATTTAGAAACAACAATCTCTGACTTTTGACCGT	1094
Db	1423	GATGCCCGCGTTTCATTTAATTAATAAACCAAAATTCGGCGAACAACAGCTCTGATTTGATCGT	1482
Qy	1095	GAAAACTACAAGAAGTTTGGGAAATTTAGCTGCTGGTGTAGCTGTTTATCAAAAGTAGGA	1154
Db	1483	GAAAAATTACAGAAGCTTTAGCTAAATTTAGCTGGCGGGTTGCTGCTTTAAAGTCGGT	1542
Qy	1155	GCTCCACACAGACAGCTTTAAAAGAAATGAACTTCGCATTTGAGGATGCTCTAAATGCT	1214
Db	1543	GCTCCAACTGAACAGAAATTAAGAAATTTAAATTTACGAATTTGAAGTGCATTAACGCCA	1602
Qy	1215	ACAGTGCAGCCGTTGAAGAAGTATTCGTTGCTGGTGGTGAACAGACACTTATTTACGGTT	1274
Db	1603	ACAGTGCAGCGCTGTAGAAGAAGGCTATGTTTCTGGTGGTGTACCGCACTGTGCAATGTA	1662

RESULT 12

RESOLUT 12
US-60-046-653-1063

: Sequence 1063, Application US/60046653

; GENERAL INFORMATION:

APPLICANT: Laqace, Robert E.

APPLICANT: Corley, Neil C.

APPLICANT: RUSSO, Frank D.

APPLICANT: Hann, Amy L.

APPLICANT: Heath, Joe D.

; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF ENTEROCOCCUS FAECALIS

; TITLE OF INVENTION: GENOME, FRAGMENTS THEREOF, AND USES THEREOF

; NUMBER OF SEQUENCES: 1

; CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACI

STREET: 3174 PO

CITY: PALO ALTO
STATE: CALIFORNIA

STATE: CALIF.
COUNTRY: U.S.A.

COUNTRY: U:
7TR: 94304

; ZIP: 94304
; COMPUTER READABLE FORM:

```

:
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk

```

```

; MEDIUM TYPE: FLOPPY disk
; COMPUTER: IBM PC compatible

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word perfect 6.1 for Windows/MS-DOS 6.2

: SOFTWARE: WORD PERFECT
: CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/60/046.653

; AFFILIATION NUMBER: 03
 ; FILING DATE: HEREWITH

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: CERRONE, MICHAEL C

REGISTRATION NUMBER: 39,132

REFERENCE/DOCKET NUMBER: PM-0003-1 P

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 855-0555

; TELEFAX: (415) 845-416

; INFORMATION FOR SEQ ID NO: 1063:

; SEQUENCE CHARACTERISTICS:

```
; LENGTH: 8784 base pairs
```

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: Linear

; MOLECULE TYPE: CDNA
; IMMEDIATE SOURCE:

```

; IMMEDIATE SOURCE:
; CLONE: FEAT1G10

```

CLONE: EFALCI063
; IIS-60-046-653-1063

05-60-046-653-1063

Query Match	56.9%;	Score 945.2;	DB 29;	Length 8784;
Best Local Similarity	74.3%;	Pred. No. 4.6e-244;		
Matches 1193;	Conservative 0;	Mismatches 413;	Indels 0;	Gaps
QY	15	ATGCAAAAGAAATCAAAATTTTTCAGCAGATGCGCGTGTGCCATGGTCGGGGAGTTGAT	74	
Db	403	ATGCAAAAGAGATTAAATTTGCGAAGATGACCTGCAGCAATGCTACGGCGAGTAGAT	462	
QY	75	ATGTTAGCAGATCCGTCAAAGTAAACGCTTGGTCCTAAAGGGCGCAATGTTGTTCTTGAA	134	
Db	463	GTATTACGATACAGTGAAGTGACATTAGCCCTTAAGGTCGTAAAGCTGTTGTTTGA	522	
QY	135	AAAGCTTTTGGTTCCTCCCTTAATTAATACGCGGGTAACCAATGCTTAAAGAGATCGAA	194	
Db	523	AAATCATTTGGTTCACCATTTAGTTACTACGATGGGTAAACAATTGCCTAAGGAAATTGAA	582	
QY	195	TTAGAAGATCATTTTGAACACATGGGAGCAAAATTTGGTGTCTGAAGTGGCTTCTAAAC	254	
Db	583	TTAGAAGATCATTTTGAACACATGGGCGCAAAATTTAGTTTCAAGATTTGCTTCTAAAC	642	
QY	255	AATGATATTGCTGGTATGGGAGACTACTGCAACAGATTTTGCACAAGCCATTGTTTCAT	314	
Db	643	AATGATATTGCTGGTGCAGCAACACAGCGACTGTTTGGACACAAGCCATTGTTTCGT	702	
QY	315	GAAGGACTAAAAATGTGACAGAGGTGTATCCAAATTTGATATCCGTCGAGGCAATTGAA	374	
Db	703	GAAGGCTTAAAAAACGCTAACTGTGGAGCTAACCCATTAGGTATTTCGCGCTGGGATGAA	762	
QY	375	ACAGCAACACACAGCTGTTGAAGCCTTGAAGCCATTGCTCAACCTGTATCTGGCAAG	434	
Db	763	TTAGCAACAAACACAGCTAGCAAGAAATTTACACAATATTTTCATCTGTAGTTGATTC	822	
QY	435	GAAGCTATTGCTCAGGTGCGTCGAGTATCATCACGCTCTGAAAAGTTGGAGAGTATATC	494	
Db	823	GAAGCGATTGCCAAGTCGCTGCTTTCATCAGGTTCTGAANAAGTCGCGCAATTAATT	882	
QY	495	TCAGAAGCTATGAGGCGTGTGGGCAACGATGTTGTGATTACCATCGAAGAAATCTCGAG	554	
Db	883	GCCGATGCAATGCAAAAAGTTGTAACGACGCGCTAATTTACCATTGAAGAAATCAAA	942	
QY	555	ATGGAACACAGACTTCAAGTGGTTGAGGCAATGCAATTTGACCGTGGTTACCTGCTCAA	614	
Db	943	ATTGAACACGAATTTAGATGTTGTAAGGAATTCGAATTCGACCGCGGTATTATCTCAA	1002	
QY	615	TACATGGTCAACAGCAATGAAAATATGGTTGCAGACCTTTGAAAACCCATTTATCTTA	674	
Db	1003	TACATGGTTACTGACAAACGATAAATGAGAGCTGTTTAGAANAATCCATATATCTTA	1062	
QY	675	ACGGATAAAAAGTGTCAACATCCAAACATTTTGGCCACTCTGAGGAAGTTCTTAAA	734	
Db	1063	ACCGACAAAAAATCTCAAATATTCAGATATCTTACCTTTATTAGAACAAATTTCTACA	1122	
QY	735	ACNACCGTCCATTACTCATTTATTCAGATGATGTGATGGTGAAGCNCCTTCCAAAC	794	
Db	1123	CAAAGCGGTCCACTATTGTAATTTGCGGATGATGTGATGGGAAGCTCTACCAACATTA	1182	
QY	795	GTCTTGAACAAGATTGCTGGTACTTTCAATGTGGTGTGCTGCAAGGCGCAGGATTTGG	854	
Db	1183	GTATTGNACAAAATCCGTGGTACATTTAATGTGTGCGCAGTAAAGGCGCCAGGATTTGG	1242	
QY	855	GATCGTCGTAAAGCTATGCTTTGAAGACATTTGCTATCTTTGACAGGTGGTACAGTATT	914	
Db	1243	GACCGCGCAAGCGATGCTTGAAGATATTGCTATTTTAAACAGGTGGTACAGTAATCA	1302	
QY	915	GAGGATCTAGGACTTGAATTTAAAGATGCTACATGACAGCCCTTGGACAGGCTGCTA	974	
Db	1303	GAGCACTTAGGGTTAGAGTTAAAGACACAACTATTGTAAAACTTTAGGAATATGCTA	1362	
QY	975	ATTACAGTTGATAAGATAGCACATTAATGTTGAAGCTTCAGGAGTTTCAGAGCTATT	1034	
Db	1363	GTAGTTGTGCAAAAGATAACACAACAAATTTGCGAAGGTGCTGTTTCAAAAGAGCA	1422	
QY	1035	GCTAACCGTATTGCACTGATTAAATCGCAATTTAGAAACAAACAACTCTGACTTTG	1094	

```

:
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (301) 309-8504
: TELEFAX: (301) 309-8512
: INFORMATION FOR SEQ ID NO: 42:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3625 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: US-09-070-927-42

```

```
Query Match          56.7%; Score 942.6; DB 14; Length 3625;
Best Local Similarity 74.2%; Pred. No. 1.6e-243;
Matches 1191; Conservative 0; Mismatches 415; Indels 0; Gaps 0;
```

Qy	15	ATGCGAAAGAAATCAAAATTTTCAGACGATGCGCGTGTGCCATGGTCGCGGAGTTGAT	74
Db	364	ATGCGAAAGAGNTTAAATTTTGAGAAGATGCACGTGCAGCAATGCTACGCGGAGTAGAT	423
Qy	75	ATGTTAGCAGATACCGTCAAGCTTAAGCTTGGTCTCTAAAGGCGCAATGTGTCTTGAA	134
Db	424	GTATTTAGCAGATACAGTGAAGATGACATTAGGCCCTTAAGGTCGTAACGTTGTTTTGA	483
Qy	135	AAAGCTTTTGGTTCTCCCTTAATTTACTAATACGGGGTAACCATTCCTAAAGACATCGAA	194
Db	484	AAATCATTTGGTTCACCATTGATTACTACGATGAGTAACAATGCTTAAGAAATGAA	543
Qy	195	TTAGAGATCATTTTGAACAATGGAGCAAAATTTGGTGTCTGAAGTGGCTTCTAAAC	254
Db	544	TTGSAAGATCATTTTGAACAATGGGCGCAAAATTTAGTTTCAGAAAGTTGCTTCTAAAC	603
Qy	255	AATGATATTGCTGGTGTGGGACGACTACTGCAACAGATTTTGACACAGCCATGTGTCAT	314
Db	604	AATGATATTGCTGGTGAAGCAACACAGCGACTGTTTGTGACAGCCATTTGTCGT	663
Qy	315	GAAGGACTAAAAAATGTGACAGAGTGTCTAATCCAATTTGGTATCTCCGTGAGGCAAT	374
Db	664	GAAGGCTTAAAAAACCTTAATGCTGGAGCAACCCATAGGTATTTCGTGGGATGAA	723
Qy	375	ACAGCACAGCAACAGCTGTGAAGCCTTGAAGCCATTCGTCAACCTGTATCTGGCAG	434
Db	724	TTAGCAACAAAAACAGAGTAGAGAAATTTACAAATATTTTCACTGTGTAGTTGATTC	783
Qy	435	GAAGCTATTGCTCAGTTCGCTCGAGTATCATCACGCTCTGAAAAAGTTGGAGAGTAT	494
Db	784	GAAGCGATTGCACAACTCGCTGCTTTTCATCAGGTTCTGAAAAAGTCGGCCAAAT	843
Qy	495	TCAGAACCTATGAGCGGTGTGGCAACGATGCTGTGATTTACCATCGAAGAAATCTCG	554
Db	844	GCGATGCAATGAAAAAGTTGTAACGACGCGCTAATTACCATTGAAGAAATCAAAAG	903
Qy	555	ATGGAACAGAACTTGAAGTGGTTGAAGCATGCAATTTGACCGTGGTACGTCGTCAA	614
Db	904	ATTGAACAGAAATTCAGATGCTGTGAAGGAATGCAATTCGACCGGGTAATTATCT	963
Qy	615	TACATGCTCAGACAAATGAAAAATGTTGCGACACCTTGAAACCCATTTATCTTAAT	674
Db	964	TACATGGTTACTGACACGATAAAATGGAAGCTGTTTGAAGAAATPCCATATATCT	1023
Qy	675	ACGGATAAAAAAGTGTCAACATCCAAGACATTTTGGCCACTACTGTGAGGAATGCT	734
Db	1024	ACCGACAAAAAATCTCAAAATTTCAAGATATCTTACCTTTATTAGACAAATTTCT	1083
Qy	735	ACCAACGCTCCATTACTCATTTATTCAGATGATGTGGATGTGAGACACTTCCAAAC	794
Db	1084	CAAAGCGTCCACTATTGATTTATTCGGGATGATGTGATGGGGAAGCTCTACCAACA	1143
Qy	795	GTCTTGACAAGATTCGTGGTACTTTCATGTTGGTGTGCTGTCAAGCGCCAGGATTT	854
Db	1144	GTATTTGAACAAAAATCCGTGGTACATTTAATGTGTCCGACGATAAAGCGCCAG	1203
Qy	855	GATCGTCGTAAAGCTATGCTTGAAGACATTTGCTATCTTGACAGGTGGTACAGTGA	914

RESULT 14

```

14
RES001
US-09-070-927A-42
: Application US/09070927A
: Sequence 42,
: GENERAL INFORMATION:
: APPLICANT: Charlès A. Kunsch
: Patrick J. Dillon
: Steven Barash
: TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
: NUMBER OF SEQUENCES: 982
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue
: CITY: Rockville
: STATE: Maryland
: COUNTRY: USA
: ZIP: 20850
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
: COMPUTER: HP Vectra 486/33
: OPERATING SYSTEM: MSDOS version 6.2
: SOFTWARE: ASCII Text

```



```

1 CURRENT APPLICATION DATA:
2 APPLICATION NUMBER: US/09/070,927A
3 FILING DATE: 04-May-2000
4 CLASSIFICATION: <Unknown>
5
6 PRIOR APPLICATION DATA:
7 APPLICATION NUMBER: 60/046,655
8 FILING DATE: 1997-05-16
9 APPLICATION NUMBER: 60/044,031
10 FILING DATE: 1997-05-06
11 APPLICATION NUMBER: 60/066,009
12 FILING DATE: 1997-11-14
13 ATTORNEY/AGENT INFORMATION:
14 NAME: Kenley K. Hoover
15 REGISTRATION NUMBER: 40,302
16 REFERENCE/DOCKET NUMBER: PB369
17 TELECOMMUNICATION INFORMATION:
18 TELEPHONE: (301) 309-8504
19 TELEFAX: (301) 309-8512
20 INFORMATION FOR SEQ ID NO: 42:
21 SEQUENCE CHARACTERISTICS:
22 LENGTH: 3625 base pairs
23 TYPE: nucleic acid
24 STRANDEDNESS: double
25 TOPOLOGY: linear
26
27 SEQUENCE DESCRIPTION: SEQ ID NO: 42:
28 US-09-070-927A-42

```

Query Match	56.7%	Score 942.6;	DB 14;	Length 3625;
Best Local Similarity	74.2%;	Pred. No. 1.6e-243;		
Matches 1191;	Conservative 0;	Mismatches 415;	Indels 0;	Gaps
QY	15	ATGCGAAAGAAATCAAAATTTTCAGCAGATGCGCGTCTGCCATGTCGCCGAGTTGAT	74	
Db	364	ATGCGAAAGAGNTTAATTTGCAGAAGATGCACGTGCAGCAATGCTACCGGAGTAGAT	423	
QY	75	ATGTTAGCAGATACCGGTCAAGTAACGCTTGGTCCTAAAGGGCGCAATGTTGTCTTGA	134	
Db	424	GTATTAGCAGATACAGTGAAGTGACATTAGGCCCTTAAGGCTCGTAACGTTGTTTTAGAA	483	
QY	135	AAAGCTTTTGGTCTCCCTTAATTACTAATGACGGGTAAACCATTCCTAAAGAGATCGAA	194	
Db	484	ANATCATTTGGTTCACCTTGATTTACTACGATGGAGTAACTATGCTAAGGAAATTTGAA	543	
QY	195	TTAAGAGATCATTTTGAACAACATCGGACGAAATTTGGTCTCTCAAGTGGCTCTCAAAAC	254	
Db	544	TTGGAAGATCATTTTGAACAACATGGCGCAAAATTAGTTTCAGAACTTGCTTCTAAACG	603	
QY	255	AATGATATTGCTGGTATGGAGCAGTACTGCAACAGTTTTTGACACAGCCATTGTTCAAT	314	
Db	604	AATGATATTGCTGGTACGCGAACCAACACGCGACTGTTTGAACAAGCCCATTTGTTCTG	663	
QY	315	GAAGGACATAAAATGTGACAGCAGTGCTTAATCCAAATTTGGTATCCGTCGAGSCATTGAA	374	
Db	664	GAGGCTTAANAACGTAACTGCTGAGCCACCCATTAGGTATTCTCGTGGGATTGAA	723	
QY	375	ACAGCAACACAAACAGCTGTGTAAGCCTTGAAGCCATTGCTCAACCTGTATCTGGCAAG	434	
Db	724	TTAGCAACAAAACACAGCTAGAGAAGTAATACACAATATTTCATCTGTATTCAAA	783	
QY	435	GAAGCTATTGCTCAGGTGCGTGCAGTATCATACAGCTCTCAANAAGTTGGAGAGTATATC	494	
Db	784	GAAGCGATTGCACAGTCCGCTGCTTTTCATCAGTTCTGAAANAAGTCGGCCAAATTAAT	843	
QY	495	TCAGAAGCTATGAGAGGCTGTGGGCAACGATGTGTGATTACCATCGAAGAAATCTCGAGT	554	
Db	844	GCCGATGCAATGAAAAAGTTGTGAACGAGCGGCTAATTACCATTTGAAGAAATCAAAAGG	903	
QY	555	ATGGAACACAGAACTTCNAGTGGTTGAAGGCATGCCAATTTTGACCGTGGTTACCTGTCTCAA	614	
Db	904	ATTGAACACAGAAATAGATGTGGTTGAAGGAATGCAATTCAGCCGGGGTTATTATATCTCAA	963	
QY	615	TACATGGTCCACAGCAATGAAAAATTTGGTTGACAGCTTTGAAAACCCATTTATCTTAATCT	674	

RESULT 15
US-09-134-000-1154
; Sequence 1154, Application US/09134000A
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al

THIS PAGE BLANK (USPTO)